

**PLAY WITH DATA – AN EXPLORATION OF PLAY ANALYTICS
AND ITS EFFECT ON PLAYER EXPERIENCES**

A Dissertation
Presented to
The Academic Faculty

by

Ben Medler

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Digital Media
in the School of Literature, Communication and Culture

Georgia Institute of Technology
August 2012

COPYRIGHT 2012 BY BEN MEDLER

**PLAY WITH DATA – AN EXPLORATION OF PLAY ANALYTICS
AND ITS EFFECT ON PLAYER EXPERIENCES**

Approved by:

Dr. Brian Magerko, Advisor
School of Literature, Communication
and Culture
Georgia Institute of Technology

Dr. Michael Nitsche
School of Literature, Communication
and Culture
Georgia Institute of Technology

Dr. Carl DiSalvo
School of Literature, Communication
and Culture
Georgia Institute of Technology

Dr. John T. Stasko
School of Interactive Computing
Georgia Institute of Technology

Michael John
Creative Director
Electronic Arts

Date Approved: June 13, 2012

To my family, friends and everyone who has supported me over the years.

ACKNOWLEDGEMENTS

I have been lucky to have Brian Magerko as my graduate adviser for my entire graduate career starting at Michigan State University and ending at Georgia Tech. This dissertation would not have been possible without his constant guidance both relating to the work presented in this document and how to navigate the academic world as a whole. He gave me a free hand to pursue my research and I am forever grateful for his support over the years.

I am also fortunate to have had an excellent committee filled with a group of intellectuals and practitioners from many fields related to my work. John Stasko, Michael Nitsche and Carl DiSalvo have been my professors and have made my academic career a success. Michael John introduced me to how the games industry functions and this has altered my perspective on academic research. All of them have been supportive and made their mark on my work.

There are also many fellow Ph.D. comrades that helped me with my work that I must thank. Paul Clifton, Simon Ferrari, Sergio Goldenberg, Thomas Lodato, Bobby Schweizer provided both excellent conversations and alternate perspectives on my work, in addition to being dear friends. There are also many other Georgia Tech Ph.D. and Master students I call dear friends as well and made my time at Georgia Tech invigorating and spectacular. Thank you: Mariam Asad, Thomas Barnwell, Hank Blumenthal, Ka Nin Chow, Jill Coffin, Clara Fernandez-Vara, Thomas Gibes, Thomas Jenkins, Tanyoung Kim, Friedrich Kirschner, Jonathan Lukens, Hye Yeon Nam, Andrew Quitmeyer, Adam Rice, Ozge Samanci, Beth Schechter, Brian Schrank, Daniel Upton, Chih-Sung (Andy) Wu, Jichen Zhu and Alex Zook. Additionally, big thanks to my CoC gaming group for giving me a place to escape from my work: Mike Helms, Josh Jones, Maithilee Kunda and Bryan Wiltgen.

I have many thanks to give to the faculty at Georgia Tech for creating an exciting academic environment and to both Janet Murray and Ian Bogost for being wonderful graduate directors. I would also like to mention and thank the Georgia Tech staff and administrative member who have helped me over the years: Grantley Bailey, Kenya Devalia, Matthew McIntyre, Alison Nichols and Melanie Richard. Furthermore, I am grateful for the funding received from the National Science Foundation and Electronic Arts that helped me complete this dissertation.

Thank you to all of my friends I made in Georgia during my Ph.D. carrier, both named above and those outside of Georgia Tech, the friends I made while interning at Electronic Arts and my friends I had to physically leave in Michigan where all of this started. I have been very lucky over the years to have such wonderful, caring friends.

Finally, I would like to thank my family for their unwavering support and devotion to helping me pursue my passions in life. None of this would have been possible without them.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
SUMMARY	xiv
<u>PART 1 - CHAPTERS</u>	
1 INTRODUCTION	1
Defending Halo Wars	1
Play Analytics	4
The Playful Side of Data Analysis	6
The Player, The Analyst	10
Designing and Experiencing Data Analysis	19
Functional and Phenomenological Aspects of Play Analytics	21
Studying Play Analytics	37
Overview of Dissertation Chapters	41
2 DECONSTRUCTING DATA	44
Defining Data	45
Data Transition and Enrichment	48
Game Data	51
Functional Data and Data Phenomenology	56
Summary	58
3 ENVISIONING VISUALIZATION	60
Your Vision	61

Information Overload and Filter Failure	65
Visualization Domains	68
The Functional and Phenomenological Aspects of Visualization	99
Summary	112
4 ANALYZING GAME ANALYTICS	114
Development Analytics	119
Play Analytics	130
The Functional and Phenomenological Aspects of Play Analytics	139
Summary	146
<u>PART 2 - CHAPTERS</u>	
5 PLAY ANALYTIC CONTENT ANALYSIS	148
Methodology for a Play Analytic Content Analysis	148
Leaderboards	156
Player Statistics	169
Group Statistics	187
Match Statistics	196
Global Statistics	206
Maps	217
Content Generation	228
Content Database	240
Discussion	251
Summary	262
6 PLAY ANALYTIC USER STUDY	265
Methodology and Study Design	267
Analysis of Study Data	273

Discussion	297
7 CONCLUSION	303
What Is Play Analytics?	306
What Play Analytics Can Be	309
Dissertation Contributions	318
The Future of Play Analytic Research	321
APPENDIX A: PLAY ANALYTIC SYSTEM INFORMATION	325
REFERENCES	335

LIST OF TABLES

Table 6.1: User Study - Time Played	277
Table 6.2: User Study - Cases	281
Table 7.1: Analytic Tasks	310
Table A.1: Play Analytic Systems – 1 to 15	326
Table A.2: Play Analytic Systems – 16 to 30	327
Table A.3: Play Analytic Systems – 31 to 45	328
Table A.4: Play Analytic Systems – 46 to 60	329
Table A.5: Play Analytic Systems – 61 to 75	330
Table A.6: Play Analytic Systems – 76 to 81	331

LIST OF FIGURES

	Page
Figure 1.1: Halo Wars	3
Figure 1.2: Properties of Play	8
Figure 1.3: J.S. Joust Exposures	17
Figure 1.4: Platform Studies	20
Figure 3.1: Visual Data Representation	63
Figure 3.2: Starcraft 2 Mini-Map	67
Figure 3.3: Visualization Domains	69
Figure 3.4: Google Analytics	71
Figure 3.5: MRI and fMRI	74
Figure 3.6: Jason Salavon's Work	75
Figure 3.7: Minimalist Visualization	77
Figure 3.8: Tag Cloud	78
Figure 3.9: Google Book	80
Figure 3.10: Gapminder	81
Figure 3.11: Critical Visualization	84
Figure 3.12: Architecture and Justice	86
Figure 3.13: Infocanvas	87
Figure 3.14: Artifacts of the Presence Era	89
Figure 3.15: Social Collider	90
Figure 3.16: Visual Diary	92
Figure 3.17: Nicholas Felton's Yearly Reports	93
Figure 3.18: Nike+	94

Figure 3.19: Fizz	95
Figure 3.20: Halo Waypoint	96
Figure 3.21: Cyclopath	97
Figure 3.22: Spore Skeleton	98
Figure 4.1: Mochibot	120
Figure 4.2: Tableau	121
Figure 4.3: Skynet	123
Figure 4.4: Data Cracker	125
Figure 4.5: Visual AI Debugging	127
Figure 4.6: Giant Bomb	132
Figure 4.7: The Sims Exchange	133
Figure 4.8: SC2Gears	134
Figure 5.1: Single Criterion Leaderboard	158
Figure 5.2: Multiple Criteria Leaderboard	159
Figure 5.3: Limited Leaderboard	160
Figure 5.4: Friend Leaderboard	161
Figure 5.5: Group Leaderboard	162
Figure 5.6: Accumulated Statistics	171
Figure 5.7: Accumulated Statistics	172
Figure 5.8: Game Mode Statistics	173
Figure 5.9: Demographic Data	174
Figure 5.10: Load-out	175
Figure 5.11: Reputation and Ranking	176
Figure 5.12: Compare Players	176
Figure 5.13: Trophy Rooms	177

Figure 5.14: Media Sharing and Storage	178
Figure 5.15: Activity Feed	179
Figure 5.16: Friend List	179
Figure 5.17: Acknowledgement Group	188
Figure 5.18: Collection Group	189
Figure 5.19: Integrated Group	190
Figure 5.20: Support Group	191
Figure 5.21: Accumulated Report	197
Figure 5.22: Detailed Report	198
Figure 5.23: Event Timeline	199
Figure 5.24: Match Leaderboard	199
Figure 5.25: Accumulated Statistics	208
Figure 5.26: Community Collection Event	209
Figure 5.27: Global War	210
Figure 5.28: Usage Statistics	211
Figure 5.29: Game Data Maps	218
Figure 5.30: Player Data Maps	219
Figure 5.31: Global Player Data Maps	220
Figure 5.32: Game Content Sharing	230
Figure 5.33: Game Content Creation	231
Figure 5.34: Non-game Content Creation and Sharing	232
Figure 5.35: Gameplay Analysis and Sharing	234
Figure 5.36: Static Databases	242
Figure 5.37: Dynamic Databases	243
Figure 6.1: Do I Have A Right?	269

Figure 6.2: User Study – Play Analytic System Part 1	270
Figure 6.3: User Study – Play Analytic System Part 2	271
Figure 6.4: User Study - Game-related Data Part 1	283
Figure 6.5: User Study - Game-related Data Part 2	284
Figure 6.6: User Study - In-game Information Formats	286
Figure 6.7: User Study - Manage Groups	287
Figure 6.8: User Study - Spectating	288
Figure 6.9: User Study - Web Browsing and Gaming	290
Figure 6.10: User Study - Play Analytics Part 1	291
Figure 6.11: User Study - Play Analytics Part 2	292
Figure 7.1: SC2Gears	311
Figure 7.2: Spore Skeletons	315
Figure 7.3: Just Cause 2 Visualization	316
Figure 7.4: Kingdom Hearts Visualization	317
Figure 7.5: J.S. Joust Visualization	318

SUMMARY

In a time of ‘Big Data,’ ‘Personal Informatics’ and ‘Infographics’ the definitions of data visualization and data analytics are splintering rapidly. When one compares how Fortune 500 companies are using analytics to optimize their supply chains and lone individuals are visualizing their Twitter messages, we can see how multipurpose these areas are becoming. Visualization and analytics are frequently exhibited as tools for increasing efficiency and informing future decisions. At the same time, they are used to produce artworks that alter our perspectives of how data is represented and analyzed. During this time of turbulent reflection within the fields of data visualization and analytics, digital games have been going through a similar period of data metamorphosis as players are increasingly being connected and tracked through various platform systems and social networks. The amount of game-related data collected and shared today greatly exceeds that of previous gaming eras and, by utilizing the domains of data visualization and analytics, this increased access to data is poised to reshape, and continue to reshape, how players experience games.

This dissertation examines how visualization, analytics and games intersect into a domain with a fluctuating identity but has the overall goal to analyze game-related data. At this intersection exists play analytics, a blend of digital systems and data analysis methods connecting players, games and their data. Play analytic systems surround the experience of playing a game, visualizing data collected from players and act as external online hubs where players congregate. As part of this dissertation’s examination of play analytics, over eighty systems are analyzed and discussed. Additionally, a user study was

conducted to test the effects play analytic systems have on a player's gameplay behavior. Both studies are used to highlight how play analytic systems function and are experienced by players. With millions of players already using play analytics systems, this dissertation provides a chronicle of the current state of play analytics, how the design of play analytics systems may shift in the future and what it means to play with data.

CHAPTER 1

INTRODUCTION

Defending Halo Wars

In late 2010, the players of the game Halo Wars were hit with a demoralizing announcement. Microsoft decreed the gameplay statistic tracking system for Halo Wars, and the entire Halo Wars website, was going to be shut down (Figure 1.1). This meant the game would no longer keep track of gameplay statistics (also called stat tracking) like the number of wins a player earned while playing Halo Wars, even though the game itself could still be played online. Removing stat tracking was a major shock to the Halo War's player community given the game's genre.

Halo Wars is a Real Time Strategy (RTS) game based in the Halo universe, a popular brand within the world of video games. RTS games have solidified as a genre over the years and Halo Wars follows many of the typical RTS motifs. Multiplayer RTS games are played based on a match system. A match consists of multiple player opponents strategically attacking each other until a final win condition is met (for example, a condition where one opponent is the last survivor). After a match concludes, players can move onto a new match where the game resets (like a chess match). Each match is played within an enclosed map environment. Players have a top-down view of the map where they can acquire resources to buy buildings. Those buildings serve as a means to produce units, which is where the strategy aspect of an RTS games is revealed. Units are built to form armies and used to attack other player opponents, specifically an opponent's units and buildings. Players must employ strategies in order to defeat their opponents such as: traversing the map's terrain, scouting, choosing stronger units to attack weaker units and managing resources. Players also have a wide variety of choices when it comes to choosing which buildings and units they use. Certain "build orders" can become common – referring to the order in which specific buildings and units are built –

but players can always deviate from these build orders, especially if an opponent knows which build order a player is using and attempts to compensate. Understanding RTS matches, therefore, becomes more than knowing who won or lost. It becomes important to understand what strategies were used in a match too.

Stat tracking in Halo Wars provides both a means to track a player's skill level (their performance over time) and the strategies players use during matches. On the Halo Wars leaderboards, for example, both win/lose records and overall scores are used to rank players. A player's Trueskill ranking (Herbrich and Graepel, 2006), a value denoting the skill level of a player, is derived from their win/loss data too and is used for matchmaking players together (players with similar Trueskill rankings are often matched together because they are considered optimal opponents with a similar level of skill). Data from each match is also recorded and describes what strategies a player executed in the game, including: where did players acquire their resources (i.e. were the resources gathered from the map or did a building produce them), the number of special abilities player use and the type of buildings constructed are counted. A player's units are tracked even further. Units are counted and shown whether they ranked up over time (became stronger), whether they killed, or were killed by, other units. Knowing which buildings were produced, abilities used and how successful units were in battle can generally reveal the build order the player used and what strategies a player executed during the game. If Microsoft shut down the Halo Wars stat tracking service all data regarding a player's skill level and the ability to analyzing individual matches would be lost.

The reason Microsoft was shutting down stat tracking in Halo Wars was due to the low population of players playing the game. On November 22, 2010 there were just over 24,000 players playing the game online, with just over 15,000 matches played that day. Overall, 24,000 player may seem like a high number but the game had also sold over a million copies in the first six weeks after the game was released and almost two million by the end of 2010 (Halo Wars Sales Figures, 2012). Less than one percent of the players who bought the game were using the Halo Wars stat tracking system.



Figure 1.1: The Halo Wars game statistics website. Data representing players, game matches and the entire Halo Wars community are presented on the website.

After Microsoft announced the stat tracking service was shutting down the Halo Wars community began to speak up. The Halo Wars forum was ablaze with harsh comments about Microsoft's lack of support for the Halo Wars community (Cocopjojo, 2010). At the same time, many gaming blogs began running stories referencing the shutdown of Halo Wars stat tracking service (McElroy, 2010; Nicholson, 2010). Comments on these news stories had mixed reactions; some thought it was terrible news while others didn't see the point of running a system if so few people played the game. For two weeks after these new stories ran the Halo Wars community continued to try and convince Microsoft to keep the Halo Wars stat tracking system.

On December 8th, 2010 – just one week before the Halo Wars website was scheduled to be taken down – Microsoft and 343 industries (the new developer handling the development of current Halo games) responded:

“Waypoint [Microsoft’s main website for the Halo franchise] and 343 Industries were always uncomfortable with the idea of taking stat-tracking offline, even though the technical and logistical problems presented made sense. The argument that it was a small and shrinking population didn’t do much to quell the reaction of both Halo Wars fans and Halo FPS fans who were concerned about the future of other Halo titles, and we decided that the current plan of action was not in line with how we have always intended to support Halo games and the Halo franchise in general.” (Lee, 2010)

Halo Wars was sticking around after all. While some features, mainly the Halo Wars forums, would be folded into the new Halo Waypoint website, stat tracking was being maintained. In fact, at the time of this dissertation’s publication Halo Wars’ stat tracking system is still running (over a year and a half later). There are even more players playing Halo Wars today – over 40,000 in March 2012 – than there were at the end of 2010, according to the Halo Wars website.

This Halo Wars story is not about players trying to save a game. Halo Wars was never in trouble of becoming unplayable. If the shutdown had occurred players would have kept the ability to play online and play through the single-player campaign. Instead, this story is about saving the community and features that exist around a game. Players would have lost the ability to have their multiplayer matches mean something more than single wins. Without the ability to expose their stats, to show off their competitive and strategic abilities, they would have no historical record of their achievements to look back on and compare with each other. They would have lost the ability to stay connected as a community. This is what the Halo Wars community was trying to save, features that turned their game into something more than just a game.

Play Analytics

Stat tracking is not a rare phenomenon the games industry in general. Since the rise of online gaming, tracking player statistics and other related data has been on the rise (Medler, 2009b). Allowing a game to connect to online services and servers makes it

much easier to remotely capture data from gameplay, store the data for later analysis or serve the data back to players. But even without acknowledging the rise of online gaming, video games have routinely in the past tracked players locally. Every time a player earns a score in a game or creates a save file, data is being recorded. What changed as broadband began becoming widely available, and game consoles became internet-enabled machines, was tracking data remotely became a viable option for game developers. All of the data being collected locally could now be gathered en masse. Data could be tracked from all over the world and new systems for linking players together began to emerge. This meant game developers could use the data they were collecting to understand who was playing their games and, at the same time, players themselves could use the data as part of their gameplay experience. Systems like the Halo Wars stat tracking system not only collect data from players (for game developers to analyze) but allow players to analyze their own data. Analyzing game data became an additional external play experience for players.

This dissertation is devoted to understanding the external play experiences players have when analyzing game data. I have labeled the domain which encompasses the systems and services used by players to analyzing game-related data as the domain of play analytics. The name play analytics is used in order to contrast the domain with another domain called game analytics. Much like play analytics, game analytics consists of “the systems and methods used to analyze game-related data” (Medler et al. 2011). However, game analytics covers any system used to analyze game-related data by any audience. A financial analyst reviewing sales records detailing how many copies a game sold can be interpreted as an example of game analytics. Conversely, play analytics is meant to cover the systems that players use to analyze game-related data and is a sub-domain of game analytics. Players are the main audience I am studying in this dissertation and are the analysts using play analytic systems.

My choice to study players is perhaps counter-intuitive to how the term analytics is regularly expressed as a serious form of data analysis, one for making vital decisions that have serious consequences. Domains like business analytics are seen as vital to a

company's overall performance and provide predictions regarding how a business moves forward (Davenport and Harris, 2007). Similarly, game developers or publishers are the more likely candidates to benefit from game analytics because analyzing game-related data can help them design future games and reap financial rewards. Players just play games, they do not need to analyze their gameplay to have fun. However, some players do find it fun. The Halo Wars community specifically fought to keep the ability to record and analyze their gameplay. Furthermore, there exists many players/hackers who create their own tools for analyzing games (Belicza, 2010; Guillaume and Joshua, 2006a; Krush DarkGod and Urme TheLegend, 2009; Meyers, 2009). Play analytic systems appear across a number of games, platforms and player communities. Players have as much to gain and benefit from game analytics as any game developer or publisher. Perhaps play analytics doesn't need to conform to the typical definition of analytics. Instead, maybe play analytics is a playful kind of analytics.

The Playful Side of Data Analysis

In his book *Now you see it*, Few presents a list of personal traits that a data analyst should exhibit (Few, 2009, p19). Some of Few's traits reflect the common, serious tone of data analysis: being skeptical, methodical, analytical, etc. However, Few begins his list with five traits that seem separate from the other serious, traits. I argue these traits form the "playful" side to data analysis:

Playful Analyst Traits

Interested

Curious

Self-motivated

Open-minded and Flexible

Imaginative

Serious Analyst Traits

Skeptical

Aware of what's worthwhile

Methodical

Capable of spotting patterns

Analytical

Synthetical

While play, or being playful, is a notoriously difficult concept to define I am using the term in relation to two definitions that present possible ‘properties of play.’ First, Caillois in *Man, Play and Games* lays out his properties of play in a sociological pursuit to describe how culture is represented through play and games (Caillois, 2001, p9-10). Second, in the book *Play*, Brown presents a similar list of play properties but structures his arguments from a clinical perspective, having studied how play is exhibited by both animals and humans (Brown and Vaughan, 2009, p17). Figure 1.2 compares Caillois and Brown’s respective property lists identifying their similarities and differences.

Both researchers agree that play is free or voluntary, meaning players are not obligated to participate in a playful activity. Both researchers view play as having an unproductive side but one of Brown’s major arguments is that play is vital for learning and living a healthy, happy life (Brown and Vaughan, 2009, p48-51). Caillois’ list goes on to state that play is uncertain and players create new sets of rules to govern how play commences. Brown on the other hand acknowledges the uncertain, improvisational potential found in play and states that rules are not necessary. There are also loose connections between the Caillois property of make-believe and Brown’s properties of ‘diminished consciousness of self’ and ‘freedom from time’. Each point to the alternate reality play creates. Caillois acknowledges a second, make-believe reality created by play and Brown acknowledges what the player experiences, a loss of self and time. Caillois’ additionally states that play is separated from the real world, which seems to point to how play functions as a practice. Brown chooses to instead highlight a player’s desire to continue playing once started, a property that speaks again to the player’s experience.

Properties of Play

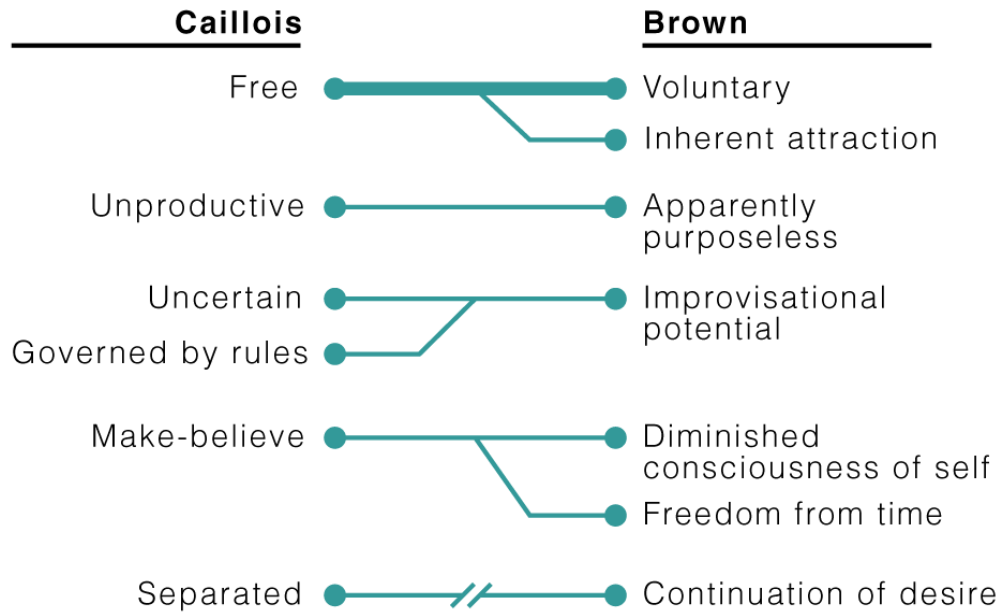


Figure 1.2: A comparison of Cailliois' and Brown's properties of play.

Each of Few's data analyst traits I label as playful can be matched with the aforementioned properties of play laid out by Caillios and Brown. Traits such as being interested and curious coincide with the inherent attraction property of play. An analyst that freely expresses interest in a data set is more likely to be able to play with that data. Self-motivated, open minded analysts set their own goals or rules creating a situation where they improvise how they analyze a data set. Analysts may even feel motivated to continue their analysis even after they have gain their initial insights from their data. Finally, analysts must be imaginative, creating make-believe scenarios and new ways to illustrate the hidden patterns in a data set.

Where Few's analyst traits and the properties of play do not concur is how play must be unproductive and separated. Analysts need a playful and a serious side to their personality according to Few's listed traits - one connects to the other. Imagining a new visual orientation for a dataset can be proceeded by further analytical and methodical approaches. These actions certainly do not need to be separated, in regards to the

‘separate’ play property Caillois mentions and happens as a natural part of the analysis process. Therefore, data analysis needs to provide both a space for analytic endeavors and a space for play too. Data analysis is both an analytical and a playful activity, an activity that is both serious and playful.

The phenomenon of moving data analysis beyond serious analytic systems has already begun within some data analysis communities, at least at the fringe. Pousman, Stasko and Mateas’ (2007) research reports a growing trend in information visualization, a type of visual data analysis, to provide a wide array of audiences with systems that do not solely focus on serious analysis. They call these types of systems “Casual Information Visualization” which they define as “the use of computer mediated tools to depict personally meaningful information in visual ways that support everyday users in both everyday work and non-work situations” (Pousman, Stasko and Mateas, 2007, p1149). They separate Casual infovis into three categories to correspond to the type of systems they covered. Ambient InfoVis are systems found in “peripheral locations and provide abstract depictions of data.” Social InfoVis systems visualize social networks and allow its users to interact with their social data. Last, Artistic InfoVis are systems with the “goal of challenging preconceptions of data and representation.”

Casual infovis systems can be seen as less productive, offer a wider variety of improvised data representations, are used on a more voluntary (i.e. casual) basis and seem purposeless to other users. Each of these descriptions is also a property used to characterize play by Caillois and Brown. Ambient Infovis systems, for example, provide displays that may include raw data streams which are less productive to interpret. Whereas a Social Infovis system may visualize data in an analytically productive way but only present data related to a single user. One user may find no purpose in interpreting another person’s data but would be inclined to analyze their own data even if only casually. Finally, Artistic Infovis systems improvise, or create their own rules for, the form and representations of data that are visualized. These are meant to express an almost make-believe view of what a data set can look like if a typical, serious data representation is discarded.

When one compares the types of systems found within casual infovis to the properties of play, casual infovis can be described as an example of how play is being introduced into infovis. Seemingly unproductive, uncertain, and voluntary systems match with how play is structured and experienced. However, the three categories of casual infovis do not denote play as a specific quality found within those infovis systems. Ambient, social and artistic information visualizations do not necessarily need to be playful. Thus, I have argue that an additional category should be added to casual Infovis, playful infovis, one that promotes play through computer-mediated visualization (Medler and Magerko, 2011). In this dissertation, I use these past arguments to create a new domain that exists within game analytics. The new domain is play analytics and it is a domain describing the ways in which play is promoted or supported through various forms of data analysis and visualization.

The Player, The Analyst

Data analysis, as I have argued, is a method of inquiry for making decisions with “serious” consequences and is not typically viewed as being playful. Analytical, scientific, utilitarian, even critical, these are the terms that describe data analysis and rarely does the term playful enter into the conversation. The prevailing view of data analysis suggests that as a practice it is better suited for assessing security threats, monitoring business strategies or conserving energy, all areas that need to be optimized and analyzed. Instead, this dissertation builds on the arguments of those that have begun to turn away from the “serious” view of data analysis (Pousman, Stasko and Mateas, 2007; Macklin et al. 2009) and declare, quite plainly, that data analysis is a form of play demanding to be studied as such. As Brown writes “the ability to play is critical not only to being happy, but also for sustaining social relationships and being a creative, innovative person” (Brown and Vaughan, 2009, p6). If data analysis is a form of play, and therefore provide the benefits that Brown suggests, we must study how those benefits manifest themselves for the players using play analytics as part of their gameplay.

In this section, I describe the common and differentiating factors between play analytics and the other theories declaring data analysis as a form of play. Macklin's work, for instance, describes how data analysis and information visualization can be applied to gameplay. Additionally, Pousman's et al. work alludes to the fact that play can be added to the domain of information visualization to form casual information visualization. As part of casual information visualization, users are given "useful" tools (i.e. less serious or casual) and are not required to approach data analysis through utilitarian means. Both works treat the player, or user, as an analyst, something play analytics does too. These bodies of work start to diverge from my own work, however, when I define play analytics as removed from the actual gameplay experience (unlike Macklin's work) and when I argue players can use play analytics for other, non-normative data analysis methods (an opposite, 'unuseful' approach to the ideals of casual information visualization).

Players are Analysts

Play analytics covers a wide variety of systems and services. There are game database websites where players explore and analyzing game content such as: in-game items, quest locations, non-player characters, strategies, etc. Players build mapping tools that chart virtual worlds or other game spaces, using them to mark locations of relevant game information that are not marked in the actual in-game map. Other play analytic systems are semi-linked with gameplay itself and offer recommendations or other data analysis features to players in-between game sessions. Some game developers also provide large gameplay analysis websites that visualize a player's past gameplay. Play analytics is a domain filled with a vast collection of different systems using a variety of data analysis methods and it is all for the player.

The main audience of any play analytic system is always the player. Similarly to Pousman's and Macklin's work, play analytics is built to accommodate a large, diverse audience of players with mixed backgrounds and intentions. Although, other audiences may find play analytics systems meaningful too. Spectators may wish to analyze data when they are spectating games. Third-party services, like social networks catering to a game playing demographic (Fong, 2007; Radoff, J. 2006), may use play analytics as an

additional feature they provide their users. Each of these audiences, particularly the players, are treated as analysts in a play analytic system.

Developers creating play analytic systems hope their audiences become more engaged with a game's content through the analysis process. For example, a play analytic system such as a leaderboard is often created by game developers to add an extra mechanism for promoting play. Leaderboards promote competition, they force players to judge how well they are performing in the game and to think about how they may earn a higher score. Other example play analytic systems are built by players who feel a game's developer is not providing the tools they want or require. Systems like replay analyzes (Belicza, 2010) or mapping tools (Krush DarkGod and Urme TheLegend, 2009) are built by players to aid them in their pursuit of analyzing their game data when developers fail to provide those players with adequate substitutes to those player built systems. Regardless, whether play analytic systems are built by players or developers, the systems are always focused on the player as an audience.

Play Analytics are Extra-diegetic

Players regularly use data analysis during gameplay as part of their play experience. Players track resources, create new content and attend to their social in-game relationships all of which are forms of data analysis. Macklin et al.'s work on the Data Play project thrives on the fact that data analysis is a part of play and has sought to combine information visualization with gameplay in hopes of creating new methods of play using real-time data analysis (2009). In addition to combining information visualization with gameplay, as Macklin's work does, another valid research pursuit is to study how players use data analysis in real time, make snap decisions or determine how to socially perform while in an online virtual space. One way such research could be framed is by using a framework laid out by Galloway (2006) that defines in-game actions along two scales: diegetic/non-diegetic and operator/mechanic. Diegetic actions in a game are actions represented as part of the game world, moving an in-game character in the game world for example. Non-diegetic actions are actions outside of the game world, such as changing menu options. Operator actions are those actions taken by the player

versus mechanic actions which are performed by the computer. These two methods for classify actions (diegetic/non-diegetic and operator/mechanic) could be used to frame the study of how players analyze data in real-time. Similar to Macklin's work players perform diegetic actions related to data analysis like monitoring vital resources and non-diegetic data analysis actions like viewing maps (assuming the maps are not part of the diegetic game world). The operator and mechanic categories can be used to describe what type of data analysis players conduct themselves verses what automatic analysis is handled by the game. Keeping track of financial records in a city building simulation game may be an automatic analysis feature but the decisions on how to spend money falls to the player as the operator. Although, the limitation of the Galloway's framework and Macklin's Data Play project, for my purposes, is they assume all actions a player may take, or thinks about taking, occur while the game is running. Nor, do they take into account actions from other players. Even though games are filled with data, Macklin's and Galloway's theories do not leave any room for player data, or actions, to exist outside of a real-time game experience.

Play analytics is a domain which contains extra-diegetic systems built outside of gameplay. These systems externalize the analysis of game data (Medler and Magerko, 2011) as an extra-diegetic action. This means the action takes place outside of gameplay as opposed to Galloway's use of diegetic/non-diegetic actions which take place during gameplay. Play analytics promote and support play through various forms of data analysis and visualization, allowing players to review and plan their future game actions outside of real-time gameplay. For instance, a simple instance of play analytics is a high score leaderboard that appears outside of gameplay. Players are able to compare themselves against other players along different measurements and filter leaderboard values based on criteria such as geo-location or their friends list. Players use leaderboards to gauge how well they are playing, spur themselves into getting a higher score and determine which players are the best at a game (at least according to the leaderboard scores). Additionally, leaderboards often appear online, separate from the game itself. This allows players to monitor a leaderboard outside of the game and can share their scores with friends without having to be in the game itself. Sharing other game-related

gameplay data such as videos, achievements or replays are also examples of how players share data externally and can be analyzed at the leisure of the player. Play analytics is thus about building extra-diegetic features which allow a player's data and the data from other players to be combined and analyzed outside of gameplay.

Players can Approach Analyzing Data Differently

If players are analysts and they use extra-diegetic play analytic systems, what data analysis methods do these systems afford? One approach may be the same one laid out by Pousman's et al. in their casual information visualization work. They describe how users can approach data analysis casually, attempting to find data "useful" instead of focusing on the "utilitarian" aspects of data analysis. A utilitarian approach to data analysis means applying analytic reasoning to the analysis process. Analytic reasoning is used to discover as many insights from the data and for those insights to have the greatest impact possible. As Saraiya et al. describes (2005), the best insights allow analysts to form hypotheses about the domain the data is from (e.g. the best insights from a biological dataset should lead to hypotheses in the domain of biology). This method for evaluating insights quantifies the results of the analysis process, thus making it a utilitarian approach (the ends justify the means). The final insights validate the method of analysis. Casual information visualization argues against the quantification of insights and instead says users can find a dataset "useful" even though they may not gain any hypothesis-forming insights from the data being visualized. However, I argue that both the utilitarian and casual methods described are similar and an additional, contrasting view point for approaching data analysis is possible.

Using terms like "utilitarian" or "useful" to describe the data analysis process represent normative approaches to data analysis. They provide objective means for determining if data analysis is beneficial to a user. An analogy to these normative data analysis methods can be found in the domain of ethics and is similar to the difference between the ethic disciplines of consequentialism (utilitarianism) and deontology. Both consequentialism and deontology are types of normative ethics, they determine 'ideally' how one ought to act. Consequentialism judges an ethical choice based on the

consequences of an action (the ends justify the means). Similarly, judging data analysis insights based on their level of contribution to a domain is a consequentialist (utilitarian) argument for the importance of such insights. Deontology, instead, judges the ethics of the actions taken to reach a goal and the intentions of the actor. It covers what ethical actions should be allowed to achieve a goal. Deontology relates to treating data analysis as a “useful” process because it does not focus on the end analysis goals but judges the actions a user can perform. Casual information visualization, for example, is still seen as demanding users to learn from and understand a dataset. This is a normative (deontological) argument because it implies that users must constantly learn and expand their knowledge when exploring a dataset. It is an argument that states, ideally, a user should continuously gain objective insights about a dataset, even if a dataset is only personally meaningful to the user and those insights do not lead to further hypotheses. While a difference exists when describing data analysis as either “utilitarian” or “useful” they both only provide a normative explanation for how users can approach data analysis. When discussing play, an act which can be argued as being unproductive and uncertain, a normative approach may not be the only way of describing how players can analyze data using play analytics.

Normative approaches to data analysis cannot see how contradictory, frivolous or subjective a dataset can be and yet still hold meaning to a user. One example is how the aesthetics of a dataset, or data analysis system, may be meaningful and hold value for a user. Schrank, in his dissertation work covering the domain of avant-garde video games, proposes that avant-garde video game designers challenge the medium of games in order to understand the medium.

“The avant-garde feels out what the medium can do beyond its assumed capacities, bulging it into directions that might seem blinding, idiotic, offensive, or perhaps just ulterior, if we keep an open mind.” (Schrank, 2010, p.10)

Schrank postulates one way avant-garde designs approach “feeling out” a medium is through the formal artistic process, or creating art “for its own sake”, while

appreciating the aesthetics of the work. This argument can be related with the Sartre existential concept of “pour soi”, being-for-itself, or in other words a person’s consciousness. Being-for-itself, as Sartre argues, is the ability to choose a way of being, a path of living. It is being for its own sake. Play analytics should be able to exist for its own sake as well. Players should not be required to find insights or useful information from a dataset. For instance, after a player earns some achievements while playing a game they may wish to display those achievements as a symbol of their game playing prowess. Players do not necessarily wish to gain insight from their achievement data. They do, however, wish to have a record of their past gameplay, a record they can reminisce about. Game data may be presented as frivolous or superficial, which may very well fit with an atmosphere of play. The aesthetics of a play analytic system should allow players to approach data analysis for its own sake and have the means to reminisce about their past experiences.

Another way Schrank postulates how avant-garde designers approach “feeling out” a medium is through political transformation. These avant-garde works seek to “affect, reveal, and transform power and society” (Schrank, 2010, p.21) through altering the perspectives of the audience. At first, this argument can be interpreted as a type of normative argument. The goal of gaining insight or information from useful data is to alter the knowledge of the analyst exploring the dataset. This goal, however, implies that accurate and clean information is given to the analyst so they can discover transformative insights from the data. A counter-example is to give an analyst an ambiguous dataset, or a broken system, and still attempt to transform their views. In Wilson’s work, regarding the development of ‘broken games’, he argues:

“Games that are intentionally designed to be ambiguous, abusive, broken, or otherwise ‘incomplete’ can help shift the focus from winning to a decidedly festive, collaborative, and intrinsically motivated kind of play.” (Wilson, 2011)

When games are intentionally left broken players are given a greater power to alter how the game functions. Players decide how the game should proceed instead of the

game rules. If I return to the analogy of ethics, descriptive ethics, a non-normative field of ethics, can be compared to Wilson’s approach to game design. Descriptive ethics is the study of what people think is right, and moral, instead of what they ought to think is right. Likewise, broken gameplay makes players decide the “right way” to play a game because the rules are not well formed. This can be argued as a transforming experience for the player, as Shrank argues, shifting the power away from the games rules to the rules a player wants to enforce. Play analytics can take a similar route. Perhaps a broken or incomplete analytic system may provide players with a means to create their own methods for analysis. There are already examples of play analytic systems which were created because a game developer did not provide players with a system the players wanted, or felt they needed (Krush DarkGod and Urme TheLegend, 2009; Minecraft X-Ray, 2010). These systems allow players to transform their gameplay experience by creating their own methods for analyzing game data.



Figure 1.3: A long-exposure visualization of the game J.S. Joust. While the visualization contains player movement data it is not an example of data analysis where players must gain something empirical or even useful from the visualization. Instead, it represents a non-normative approach to analytics.

Players can therefore use both normative and non-normative approaches to analyze data as part of play analytics. In order to consolidate these two types of

approaches I use the term meaningful to describe how data can relate to a player (or a user). I define meaningful as “to provide” or “to provide for”. Gaining utility from a dataset or finding data useful are ways of describing data as meaningful to a user. Similarly, saying data exists for itself, or for its own sake, and allowing users to define the value of data for themselves are also ways of describing how data is meaningful. Play analytics can sometimes be serious and utilitarian while at other times frivolous or for its own sake. Either way, players can find meaning in a play analytic system using various approaches.

Play analytics turn players into analysts and use extra-diegetic systems to explore game data. Players can also approach analyzing game data differently when using a play analytic system. They can take normative approaches which are typical to analyzing data, approaches like using analytic reasoning to find insights or patterns in a dataset. Other, non-normative approaches allow players to experience data for its own sake, to reminisce about past gameplay, or provide players with ambiguous data or tools to create their own analyzing methods. In chapter 3 I discuss these different data analysis approaches through the lens of visualization, defining five ways to approach data analysis, and review how these approaches allow an analyst (e.g. the player) to find data meaningful. In order to explore these approaches I need a framework for understanding how meaning is applied to data. On the one hand, meaning is applied to data through a data analysis system. Systems can be designed to be analytical and support an analyst in finding correlation or patterns in a dataset. A system can also be designed to be broken or uninformative too, which may be meaningful in the context of a playful situation. On the other hand, players also apply meaning to data as part of the data analysis process. Players may wish to find insights that can help them become a better player or they may use data as a way to brag about their accomplishments to their friends. The way a data analysis system is built and how the player decides to experience data alter the data analysis process and can provide meaning to data. Therefore, in order to continue to explore how play analytics, and the data analysis process, is meaningful in respect to games, I need to examine how data analysis systems are built and how they are experienced by players.

Designing and Experiencing Data Analysis

Studying play analytics from the perspectives of how analytic systems are designed and how players experience those systems can be seen as an approach similar to platform studies.

“Platform Studies investigates the relationships between the hardware and software design of computing systems and the creative works produced on those systems.” (Bogost and Montfort, 2009)

Bogost and Monfort, the creators of platform studies, argue that the platforms on which digital systems are built affect the design and experience of those systems. However, platforms are merely a single level or a way of framing digital systems like a play analytic system. Bogost and Monfort lay out a five-level hierarchy describing different levels one can take when studying new media artifacts (Figure 1.4).

The lowest level, the platform, describes the hardware a digital artifact is built upon. Next, you have the code level that includes how code is created and compiled on a specific platform. The form and function level deals with the format and functionality of a digital system. In game studies this means analyzing things like game mechanics or narrative structures. The fourth level, interface, is concerned with user interfaces and input devices. Finally, the reception and operation level covers the reactions of the audience and how they operate a digital artifact.

Play analytics can be explored as existing along this hierarchy of platform studies. The platform and code level of play analytic systems are typically web based. Many play analytic system exist online and factors such as web browser capabilities and web coding languages do factor into a system's design. The form/function and interface levels are where the design of play analytic systems resides. These levels can be used to cover how play analytic developers create systems to visualize and analyze game data. Finally, the reception/operation level can be used to describe how play analytic systems are experienced.

As I have already argued, players and other spectators are the main audiences for any play analytic system. Their experiences are the ones I wish to address. Additionally, these experiences are normally outside of the real-time gameplay thus when discussing experiences with play analytic systems these experience stand separated from the gameplay experience (there are some exceptions to the rule and these will be discussed starting in chapter four). Players can also use many different forms of data analysis methods, some normative and others non-normative. This fact symbolizes a tight coupling between the player's data analysis experience with the overall design of a play analytic system because a system may only be designed for specific data analysis methods. The purpose of studying play analytics in a similar way as the approach dictated by platform studies is to gauge how the design of play analytic systems and the experience of those systems intermingle and affect one another. In order to do this I use a framework that allows me to examine play analytics as both an aspect of functionality (design/development) and one of phenomenology (experience).

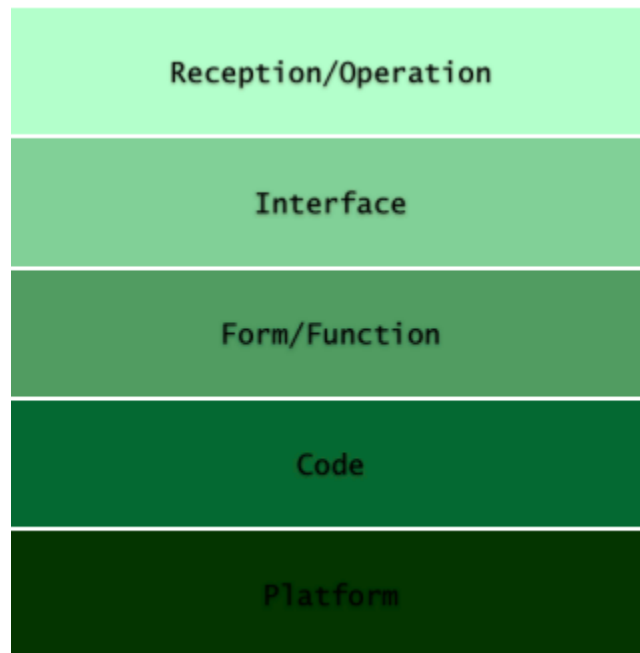


Figure 1.4: The five levels of platform studies. This dissertation studies play analytics between the 'Form/Function' and 'Reception/Operation' levels.

Functional and Phenomenological Aspects of Play Analytics

If we find play analytic systems valuable based on the functions those systems afford and we find further meaning through how we experience those systems, then I propose the process of interpreting and critiquing play analytics in this dissertation to be carried out by examining data through the lenses of function and experience.

Functionally, play analytics – and by extension data analysis as a whole – provides many benefits including gaining insights, increasing play performance, promoting competition and curating sharing or communication. Analytic systems are routinely built to cause these benefits because system designers wish to build tools for understanding data.

However, play analytics can also function poorly, be broken or ambiguous, and still be designed in such a way as to provide meaning to players. Systems can be designed to give players more freedom to add their own data or create their own analytic systems outside the control of a game's developers. Interpreting play analytic systems functionally means understanding the features or attributes, even attributes that are seemingly ulterior to data analysis, for defining how a system is meant to add meaning to data. Phenomenologically, play analytic systems are digital artifacts players experience and provide meaning to the world around them. The functionality of a system gives players a starting point for determining how a system can operate as part of their living experience but it is merely a starting point. When faced with using and living with an artifact, whether mechanical or hand-crafted, people forge their own experiences with the artifact. The theories of participatory culture argue as much (Jenkins et al., 2006), people attach their own contributions onto artifacts and other cultural entities. Interpreting play analytics phenomenological means describing and monitoring how players themselves alter the meaning of data and the analytic systems they are provided. Taking these two methods of interpretation, functionality and phenomenology, I present the theoretical framework I use throughout this dissertation to explore play analytics.

The Function of Play Analytics

Data analysis tools meant to aid in gaining insights or increasing efficiency typically are maintained as some concrete system. There are pieces and procedures that must be developed and connected in order to provide users with the ability to gain

meaning from data. Various researchers and practitioners have expressed their views of what pieces combine to create a data analysis system. For example, Fry argues that data analysis requires many processes: acquire and parse data, filter and mine data, represent and refine data, and interact with data (2004). Spence defines information visualization, i.e. using visuals to aid in the data analysis process, can be defined as having three processes: selection of data, encoding of data, and presentation of data (2000). Even though Fry lists more processes compared to Spence's list, they are fundamentally mentioning the same requirements for conducting data analysis. Acquire/parse are equivalent to selection, filter/mine/represent/refine are types of encoding and interaction can be seen as a type of presentation. Other work by Thomas and Cook (2005) also make similar claims regarding the required pieces of the data analysis process, namely: data gathering, visualization, analysis and dissemination. These pieces are almost identical to Fry's and Spence's lists. Additionally, Card et al. (1999) defining information visualization using five attributes: computer-supported, interactive, visual representation, abstract data and amplifying cognition. Attributes such as computer-supported and amplify cognition do not appear in the other lists but relate to other attributes of data analysis. Computer-supported points to a likely platform a data analysis system may be built on and the abilities of that platform to visualize and present data. Amplify cognition is the result of building a system that allows data to be analyzed; Filtering, mining or encoding can be examples of ways cognition can be amplified if we refer to the other list defining data analysis. All of these definitions list the concrete components that must exist as part of the data analysis process.

Combining the data analysis lists may get us a super list that contains such pieces like: platform, data gathering, data transformation, visualization, analysis, presentation. However, one key point is not mentioned in any of the list discussed above or in this combined list. Someone must control all of these pieces when developing a data analysis system. Someone must define what data is gathered, how the visualizations are structured and what interactions are allowed. Fry does mention that disciplines like computer science, mathematics/statistics, graphic design and human-computer interaction can be related to the different pieces of the data analysis process. Computer science skills are

used to gather data, mathematic skills are used to transform/encode data, graphic design skills are used to construct the visualization and HCI skills are used to build interactions. These relations only elude to the fact that certain pieces of the data analysis process lend themselves to certain skillsets. They do not necessarily discuss who has control over each of the pieces. A single person may possess all the skillsets mentioned and therefore have unabated control over how a data analysis tool is developed. Furthermore, these lists also do not acknowledge how a data analysis may hold control over the audience of a data analysis tool. If a tool is built to only collect a certain type of data, only displays certain visualization, only correlates data in a specific way these are factors that control how a user may approach using the tool. Users may find way to circumvent this control, which is a topic for how users experience data analysis, but, nonetheless, the controls exist as part of the development of the tool. As Derrida writes, “the science of [data] must include the theory of this institutionalization, that is to say, the theory both of law which begins by inscribing itself there and of the right which authorizes it” (where data is substituted for archive) (1998, p.4). In addition to all of the pieces that have been used to define data analysis, an element of control must be added as one of the pieces to the list.

In order to combine the lists defining data analysis with the concept of control I turn to the work of Bowker and Star. As part of their larger body of work studying how people inter-operate using data, Bowker and Star cover how categorization systems are created. In their book *Sorting Things Out*, Bowker and Star set about exploring the social implications of categorization systems which are systems setup to organize, control and institutionalize data. They cover example categorization systems like the International Classification of Diseases (ICD) which they describe as “a panoply of tangled and crisscrossing classification schemes held together by an increasingly harassed and sprawling international public health bureaucracy” (Bowker and Star, 2000, p21). Societal factors such as how work-related diseases are recorded for insurances purposes or how moral medical issues are handled (when life begins for example) contribute to the tangled system that has been evolving for decades. Bowker and Star also look at the categorization system maintained in South Africa during apartheid that classified and reclassified citizens into different racial groups, calling out problems with the “moral

accountability [of classification] in the face of modern bureaucracy” (Bowker and Star, 2000, p.196). They also cover the creation of the Nursing Intervention Classification system that is meant to not only provide the means of scientifically studying nursing intervention but to “defend the professional autonomy of nursing” as well. What all of these classification systems have in common is they can be described almost identically to a data analysis system. Each of them gather data, transform/encode data, offer methods to analyze and represent (e.g. visualize) data, disseminate data and are system which exert control, particularly over the audiences using (or being forced to use) the categorization system. Studying classification therefore gets to the heart of how standards and organization are applied to data analysis (by defining data, how data is represented, visualized, etc.) and how control is exerted throughout the system. This makes Bowker and Star’s theories of classification a proper basis to study and interpret how play analytic systems are designed and developed.

Bowker and Star postulate there are three properties of a classification system: visibility, comparability and control. Visibility describes how and what data is represented. In comparison to the lists defining data analysis, visibility represents how data is gathered, transformed and, to some extent, visualized or presented. Visibility is the property through which a categorization system can turn meaningless data into meaningful data. Data is made visible by defining the rules for representing data, allowing data to be referred to and for users to become aware of the data. Bowker and Star give the example of the direct and indirect care nurses provide patients as an example of how visible and invisible data work. If indirect care provided to a patient, e.g. constant monitoring of their vitals, is not classified then that time spent on indirect care can be seen by hospital administrators as “fiscally wasted”. Conversely, direct care such as administering anesthesia (Bowker and Star, 2000, p.245) is likely to be classified and easier to acknowledge as vital to patient care. Anything invisible to a system cannot be recorded and monitored, and therefore verified. Play analytic systems, similarly, make specific data from games visible and attempt to make players more aware of what is being recorded. Data points such as score or kill/death ratios (the number of times a player killed in a game versus how many other players they kill) are shown to players as

a means of evaluating how well they are performing in the game. Other actions may be invisible in these systems such as how often I assist other players. Games have the luxury of being simulation with a limited set of rules and actions compared to the real world, and it makes it easier to highlight those rules and actions in data. If the point of the game Pac-man is to eat pellets then it is an easy decision to record and make visible how many pellets are consumed. As play analytic systems grow, however, incorporating more features involving player socializing and player-content generation it is becoming harder to decide what data points should be made visible over others, especially if players are playing the same game for different reasons.

Once data is made visible the next step is the compare data points together. Comparability is how data is related or compared within a system in order to create information and is important “so each actor may fit their allotted position in a standardized system and comparisons may be communicated across sites” (Bowker and Star, 2000, p.241). Visualization, analysis and amplify cognition are pieces of data analysis which relate to comparability. It allows for normalization, to allow disparate systems to compare themselves along a single system. For places such as hospitals which each have their own local idiosyncrasies, comparability makes it so one hospital can be compared with another, without having to relate the idiosyncrasies for every hospital pair.

Games compare players all the time. It is the basis for most form of competition. My score is higher than your score based on a standardized system for accruing points because I knocked down more pins or collected more coins. Although, comparison can be used to compare oneself with oneself as well. EA Sports Active for instance can track how many calories a player has lost over time, providing a calendar to compare their workout results over time. Comparison can be investigative too. Correlating data, searching for patterns and trends, is an important part of many analytic tools. Systems such as SC2Gears allow Starcraft 2 players to analyze their data across multiple play sessions and compares common data points such as how fast a player is at performing actions in the game or what order structures in the game are built. Finally, comparison may be non-normative and provide no added value to gameplay itself. The system

entitled Spore Skeleton (Meyers, 2009) visualizes the skeletal structures of creatures made in the game Spore. It is a system which gathers, transforms and visualizes data from the game Spore (even to the point someone could use the tool for analysis) but the purpose is not to compete or attempt to gain anything useful to how I play the game. It is more art piece than analytic tool but still follows a similar design and structure. Many methods for comparing data are used in play analytic systems but it is important when studying these systems to identify what data is being compared and for what purpose.

Finally, the control property covers the people, procedures and institutions governing how a categorization system functions and what data it controls. Developers of categorization systems dictate the reasons, and the means to change, why data is made visible and compared. They are also susceptible to their own prejudices or interests which affect how a system is used. For example, hospital administrators may be interested in classifying nursing work, according to Bowker and Star, for the purposes of making budgets and organizing schedules. But focusing on building a system with only the thought of budgets in mind may result in nurses falling under further scrutiny or surveillance, affecting how they perform their job.

In a game, rules automatically control what a player can do and achieve. Play analytics is really a system for controlling the representation of data on top of a system (the game itself) built to control players. A first person shooter game is therefore going to be more likely to have a play analytic system that compares data surrounding the act of shooting and the competition between players. It is certainly possible to design play analytics so that more control over how they are used falls to the audience, as I mentioned when discussing Wilson's concept of broken games. That fact brings up another factor, who is in control of a play analytic system? Game developers are by far the majority group who end up controlling analytic systems for their games but this isn't always the case. Third-party services and players sometimes maintain their own systems outside the scope of what game developers provide. Studying both who controls a play analytic system and the reasons for exerting control over a system are two factors to be scrutinized when studying play analytic systems.

I argue Bowker and Star's three properties of categorization systems can be expanded to interpret any system which handles data, including play analytic systems. As I have shown play analytic systems can be described as comparing data made visible to players and exert various levels of control over those players. They can be used to analyze player performance, promote competition, allow players to share data, and create artistic pieces, all of which point to the different ways players can interpret data. But as Bowker and Star's analysis proves it is not enough to study categorization systems (or play analytic systems) based on their functional aspects alone. People must use and experience these systems. For that I turn to theories found in phenomenology and explore what it means to experience data from our own point of view.

The Experience of Play Analytics

As the examples from Bowker and Star's work covered, using and experiencing categorization, or data analysis, systems impact people's lives. Hospital workers are forced to categorize their daily routines in a similar way player gameplay is categorized in a play analytic system. However, local behavior and idiosyncrasies regularly appear within any categorization system (Bowker and Star, 2000) and play analytics is no different. No matter what system is being used, people forge their own experiences regardless of the system's overall design. It is the opposite side of functionality; people interpreting, altering and molding a system to create personal experiences.

In order to interpret how a players experience play analytics I rely on the discipline of phenomenology, the study of how we are conscious of our environment and the experiences we perceive. Using phenomenology to approach studying player experiences, I rely on a few key phenomenologists and explore how their theories of experience relate to data and play analytics. These philosophers (Husserl, Heidegger and Gadamer) and their theories were chosen for three reasons. First, phenomenology allows me to apply anecdotal evidence as a tool for describing the experience I personally have with play analytic systems. As part of this dissertation, I cover (and used) over 80 different play analytic systems. Being able to describe my own thoughts about how I

experienced these systems provide a consistent interpretation throughout the entire set of systems compared to collecting other player's impressions of individual systems. My single viewpoint creates an obvious skew within my descriptions but viewpoints collected from other individuals would certainly be skewed as well. For example, other players may only use a few play analytic systems or enjoy a system specifically because of their fondness for the game the system is attached to. Recording my personal experiences across all the play analytic systems I cover allows me to form hypotheses regarding how play analytic systems exist as a whole, because I can relate systems to one another. These hypotheses may provide future directions for studying and interpreting play analytics; directions that can include further study beyond my single viewpoint.

Second, the phenomenological theories I use are based around experiencing objects and data being interpreted as an object; although, phenomenology is not limited to the experience of objects alone. Another common concept often found in phenomenology is embodiment. The concepts of the human body and space, in relation to experience, are much more prevalent when discussing embodiment. Works from philosophers like Merleau-Ponty (1945) and de Certeau (1984) often discuss the body's role in our conscious experience. For example, de Certeau in his book *The Practice of Everyday Life* (1984) talks about how city dwellers often repurpose and circumvent their urban surroundings such as taking shortcuts or congregating in areas normally not designated for socializing. A city dweller's movement throughout the city becomes the basis for their overall experience of the city. Likewise, Dourish's work in *Where the Action Is* (2004) applies phenomenological theories surrounding embodiment and uses them as a bases to explore the field of human computer interaction and how users interact with interfaces. A phenomenological interpretation of play analytics focusing on embodiment would certainly be a valid approach, especially when compared to the work of Dourish. However, I argue play analytic systems represent an experience where players are playing with data, meaning they are playing with an object. Many of the interfaces used by play analytic systems exist in a screen-based format where players can manipulate and analyze the data. A player's body does not necessarily factor into how a player experiences the data being presented to them. Thus, I leave the application of

embodiment to play analytics as a future endeavor and instead focus on interpreting our experiences with objects.

Third, the phenomenological theorists I have decided upon allow me to synchronize their arguments with the properties used to describe categorization systems. Visibility, comparability and control each have a counter property related to how players experience a play analytic system. I leave the detailed explanation describing how and why these properties relate to each other for the rest of this section but the three phenomenological properties are: exposure, interconnection and historic. These three properties respectively connect to the three properties visibility, comparability and control. Exposure represents the experience of data being exposed to other people after data is made visible. Interconnection relates to comparability as the experience players have when comparing game data or comparing themselves to other players. Finally, the historic property encompasses how control occurs over time. Control is never finite or merely present, it evolves over time and affects how the experience of a system progresses.

Phenomenology allows me to describe my own experiences, to show how we experience objects and relate theories of experience to the functional properties defining categorization systems. Below I step through the phenomenological theories I have chosen. Each set of theories from the three philosophers I discuss (Husserl, Heidegger and Gadamer) are also related to the three properties I argue as describing the experience a player has when using a play analytic system (exposure, interconnection and historic). The three phenomenological properties are also components of my main dissertation thesis, which I cover in the next section of this chapter.

Exposure - Husserl: Bracketing and the Ideal Experience

When we look at an object what do we see? We see its color, its shape, its size, or at least that is how Husserl believes we look at objects. His theories detailing object experience include such methods as bracketing, reduction and retention (Husserl, 1963). Bracketing, in particular, is used to focus our conscious experience upon the objects we

encounter. By bracketing an object, I can think about the object objectively, disregarding other factors about the object and solely focus on the object itself. When I focus on a car I can see its color, the shape of the car's body and the size of the tires. I begin to see an object that I can definitely say is a car. I achieve this by reducing the object to its separate parts (color, body, components) and find an ideal structure I can recognize (Husserl calls this noema). I retain those structures in my memory and can retrieve them later when I encounter another object with similar properties. Husserl, thus, is speaking about how we can have an ideal experience of an object, an objective experience where we reflect on the essential structures of the object.

While the next phenomenological theories I lay out dispute Husserl's claims of bracketing and experiencing the ideal object, these two claims can be used to practically describe design. When designers and developers are building systems involving data – like a categorization or play analytic system - they have to bracket their assumptions and learn to focus on what is important to their system. They have to decide what data is important enough to provide users and how the system will function as a cohesive entity. As authors of a system, designers place with own intentions and ideas into what the ideal structures of the system should reflect. At the same time designers also hope that their users will share their ideal interpretation of the system too, and use it in the ways the designers built the system to be used for.

Bracketing and the ideal experience create a sense of exposure for players within a play analytic system. When designers build a play analytic system they create a subset of data to represent a game or a player. In the context of bracketing, the subset of data lacks connection to the outside world and is meant to be used within the confines of the system. However, data is often used outside of the arena where it was initially captured and used. What is seemingly private data becomes public or public data when a specific system is migrated to another. Acknowledging these outside factors affect the experience of an object begins to nullify Husserl's theory of bracketing but the theory can still be useful to describe why a system intentionally exposes data. Bracketing allows me to interpret the data made visible by designers as there way of creating an ideal experience

by creating a set of exposed data for players to use. The data must be exposed in order to create the ideal experience. I cannot create a system to analyze someone's gameplay without exposing that data. Although, after the data has been exposed, the designers do not have control over how the data is used. Designers expose data to create an ideal experience but there is a chance the data may not be experienced in that ideal way, hence why I use the term expose (since it is vulnerable).

Interconnection - Heidegger: Contextual Interpretation and Moods

Husserl's argument detailing how we experience objects within our environment (Husserl also uses the term life-world) describes the experience as a very precise activity, one where we determine the objective properties of an object. Heidegger takes a different approach. He instead views our experience of objects as mundane most of the time and rarely includes an object's specific physical traits. For instance, say you are looking at a parked car. From Husserl's point of view you see the car's shape, its size, its color and any number of other physical attributes which makes the car an entity in our environment. But when was the last time you looked at a car and began thinking about its physical features? Most likely when you look at a car you are thinking about your destination, the need to purchase more gasoline or, if you are with someone, may not think about the car whatsoever. Heidegger argues this nonchalant attitude towards objects is how we experience our "world" or environment. He describes it as the "average everydayness" of "being-in-the-world" (Heidegger, 1962, p.69,80. We are more likely to see objects as affordances – where will this car take me or what can this table hold – rather than having physical "objective" attributes. Experiencing objects thus becomes an experience of action and purpose where objects take part in our tasks or objectives. He acknowledges Husserl's method of bracketing can occur (i.e. deep reflection) but is merely a particular way of experiencing life. Normally we experience average everydayness.

Heidegger separates the "average everyday" experience of objects and Husserl's method of experiencing "ideal" objects into two concepts. First, there is "ready-to-hand" which Heidegger defines as when "objects relate to each other and constitute a realm of

‘significance’ most obviously and easily of objects of use” (1962, p.98-99). Objects are experienced within a specific context, which generally means how an object relates to its environment, and finds ‘significance’ within that context. Objects are contextually interpreted, they have significance in what they have been or will be used for. A car is placed into context by its ability to traverse large distances quickly, so we see it as a means to get us to a destination. However, we do not automatically gain an understanding of an object and its context. Only by using objects, exploring their context within our environment, can we gauge the full range of contextual capabilities. A child laying eyes on a car for the first time does not see the object as a tool for travel, only by experiencing the car do they learn the object’s context.

“Present-at-hand”, the second of Heidegger’s concepts, is used to describe a type of reflective experience one can have with an object and the concept closely aligns with Husserl’s method of bracketing. While we normally experience an object as ready-to-hand, when we reflect on the object it becomes present-at-hand. If my car breaks down, for example, it ceases to be seen as an object of travel and instead I view it as a complex machine. I am thrown into a world of engines, parts, fluids and electronics. The properties which form the car become apparent to me as I bracket the car’s existence because the object no longer operates in an “average” manner. But, as Heidegger points out, our experience of the car eventually turns back into ready-at-hand. Before the car was an object of traversal, now it is an object to be repaired. A car to repair is the object’s new “average”. I think of repair costs and car renting so that I may return to my average everyday activities. Present-at-hand is therefore a state when we are forced, more so than we may choose, to reflect on an object.

Both present-at-hand and ready-to-hand are methods of experiencing objects and this includes data. Ready-to-hand is a way of interpreting how we experience data in our average everyday lives. When I see a photo of my friend online I don’t think about the photo’s aspect ratio, I think about the last time I saw my friend or when the photo was taken. In a game with a health bar I don’t think about the exact amount of health I have at each moment of the game. Instead, I think about how much risk I can take because I have

enough health to spare. Present-at-hand, on the other ... hand, is when I have to look at data in a reflective manner. If the photo of my friend is blurred I notice the resolution. If I die in a game I attempt to determine how my previous actions removed so much of my health. As the physical context of the situation changes towards data, or as my mental frame of reference towards data shifts, I flow between ready-to-hand and present-at-hand, between reflection and action, causing shifts in my interpretation of the data around me.

I should point out Husserl and Heidegger use reflection in a manner that describes both analysis, the process of breaking up an entity to examine it, and reflection (a process of reexamination). One “reflects”, by Husserl and Heidegger’s terms, on an object in order to learn new properties (during present-at-hand) or affordances (during ready-at-hand). But we can also “reflect” on an object in relation to our past experiences, not necessarily to gain insight through examination or to find previously unknown properties (as I discussed in the exposure section above). Present-at-hand for example may be a state when we choose to reflect on the properties of an object, properties we may have already known but do not acknowledge most of the time. Ready-to-hand is mainly about reflection in that we “retain” (a word both Husserl and Heidegger use to refer to experiences remembered) past experiences that define the affordances an object is capable of and we then put those affordances to use without the need for examination. In this dissertation I use the term analytics in a way which covers many forms of data interpretation, including the way reflection is used by Husserl and Heidegger. I do so because reflect, by today’s terms, carries with it a sense of passive thought whereas analytics is seen as a process of active engagement. In order to capture both an active and passive view of data I differ to the term that represents action, i.e. analytics, if only to not trap myself to the realm of passive reflection alone. Thus, my version of analytics includes both a method of analysis and reflection.

Finally, Heidegger also uses the concept of moods to discuss how our “state-of-mind” affects our interpretation of objects in our environment (1962, p.172-179). Moods are more similar to frames, rather than a context, because frames refers to a mental state. My frame of mind, or mood, towards an object changes even if the context of the object,

the physical context, does not change. A car is still a car, complete with the affordance of traveling, but if I'm not in the mood to travel (perhaps I am tired) then I don't see the car within the context of traveling. "To be in a certain mood is to view the world in a certain way, and it crucially affects our engagement with the world and the ways in which we respond to entities within it" (Inwood, 2000, p40-41). A tired mood causes most object affordances in our world to become less significant, for example. However, moods should not be thought of as emotions, since emotion is typically a directed mental state. 'I am tired of reading this book' is an emotion, the significance of reading the book diminishes, whereas 'I'm tired in general' is a mood, the significance of all objects diminish, except perhaps for my bed. Moods shift and change, but we are always in a mood even if that mood is our "average everyday" mood.

The implications of Heidegger's theories on play analytics are twofold. First, the context of data has an impact on our interconnection to data, altering how it is perceived and used. A player's score by itself is perceived differently compared to displaying the score in relation to other players. Leaderboards for example rank player scores. In this case a player perceives their score as a rank, not merely a simple number. A leaderboard affords the ability to compare players because scores are put into the context of other players. A score can be seen as ready-to-hand or present-to-hand as well. When competing for the top spot on a leaderboard a player may only care about continually playing the game in order to increase their score (ready-to-hand). If the player continues to play but never advances on the leaderboard they may begin reflecting on their gameplay (present-at-hand) in order to determine how they may increase their skill and score. Leaderboards do not provide the player any further ability to reflect on their score, however, they only provide the final result and not an analysis of how a player can improve their score. We can therefore ask, as one example of examining play analytic system using Heidegger's concepts, how may a player analytic system be built to offer players the ability to reflect on their gameplay once they reach a present-at-hand state?

The second implication of Heidegger's theories is that understanding the player's mood can also affect how a player interconnects and approaches features in a play

analytic system. Players can be in any number of moods when it comes to using a play analytic system. A competitive mood, a relaxed mood, a hurried mood. Systems built on different platforms, such as a website versus a mobile application, may be better suited to accommodate for different moods too. A mobile version of a system may focus on quick digestible data because players may only be in the mood to quickly glance at the application when they are away from their game console or computer. Additionally, specific features of play analytic systems can be interpreted as only providing affordances to certain moods. Leaderboards benefit players in a competitive mood but other, less competitive, players may find leaderboards of little significance. What Heidegger's theories provide are ways to judge how play analytic systems connect players to one another and to data, while taking into account the context of the data and the mood of the players.

Historic - Gadamer: Historical Context and Prejudices of Interpretation

If it is not apparent, Husserl was a predecessor to Heidegger (Husserl actually taught Heidegger) which is one reason why Heidegger's concepts follow and expand upon Husserl's thinking. The final phenomenological philosopher I discuss is Gadamer who was influenced by Heidegger in turn. His work expands upon Heidegger's idea of contextual interpretation by including the argument that historical context has a major role in how we experience objects.

“The great historical realities of society and the state always have a predeterminate influence on any experience’.” (Gadamer, 1994, p.276)

Gadamer's theories regarding experiencing objects focus on how our past and cultural history affects our interpretations of present objects (Gadamer calls this the “effect” of history). We are always caught in a tug of war between our own interpretations of an object versus how history has interpreted the object, and how those past interpretations have affected us. “Long before we understand ourselves through the process of self-examination, we understand ourselves in a self-evident way in the family, society, and state in which we live” (Gadamer, 1994, p.276). Long before we understand

objects, we understand them within a historical context of our society. Gadamer cautions that we should be aware of this fact and to make sure when we interpret objects we take the historical interpretation of the object into consideration along with our own interpretations. If, for example, we look at certain car brands which have lasted for decades, such as Corvette or Mustang, it is misleading to discuss these objects solely based on their current model and position within our society. The history of such cars must be accounted for if we are to discuss topics such as why these brands are popular. We can explore games in this manner too. Older games from the early years of arcades and home console systems can be interpreted as primitive and frustrating by today's standards. But making those interpretations limits one view on how those games had a major impact on the progression of the game industry over the years. Unlike Heidegger's contextual interpretations which generally revolve around interpreting an object's place within its present context, Gadamer sees those moment to moment interpretation of an object throughout history as being strung together to form a cohesive interpretation link that must be explored. "The important thing is to be aware of one's own bias so that the text [or object] can present itself in all its otherness and thus assert its own truth against one's own fore-meanings" (Gadamer, 1994, p.269). We must allow the object's history (its 'other' interpretation) to mix with our own meanings and biases.

The biases Gadamer speaks of are the "prejudices" we have towards objects we experience. He defines prejudice not as a negative term but as "a judgment that is rendered before all the elements that determine a situation have been fully examined" (Gadamer, 1994, p.270). This does not mean prejudices are false or ignorant judgments. Since we are in constant dialog with our own interpretation of objects and the historical interpretations of the object we can only ever make prejudgments. There is never a time when we completely understand the whole truth about an object and can only attempt to understand our own prejudices, and the prejudices others have historically had, in relation to objects. Gadamer is arguing that historical objectivism, where one attempts to interpret the past through an objective eye in the present, is a figment and not possible. "To be historically means that knowledge of oneself is never complete" (Gadamer, 1994, p.302). We can never know the full story of the past because we are not living in the past, a past

which has also altered how we perceive the present. However, acknowledging our own prejudices and attempting to understand the prejudices others had in the past can aid in our analysis of the objects we experience. Understanding our current view and methodology towards analyzing games today can be combined with an attempt to understand what it was like to experience games in an earlier era, for instance. Placing objects in an historical context therefore must include an analysis of our own prejudices and those prejudices of the past.

Gadamer's historical context theories provide an interesting historical interpretation on studying play analytics. As systems such as Rockstar's Social Club or Microsoft's Xbox Live achievement system have continued to operate for many years how can we begin to interpret the effects of these systems over time? These systems are accumulating data continuously, interpreting the experience one has with that level of historic data is different from experiencing data collected only yesterday. Game data may eventually turn into a type of photo album that provides a historic look at our gameplay over time. However, a play analytic system can be shut down after a few years, as they are known to do (Mitchell, 2012). All of the accumulated data may be lost or decay into an unusable state. Even though the history of play analytics, or rather digital game analytics in general, does not stretch back far, approaching play analytics from a historical perspective will grow more complex as the domain progresses. The systems will become more complex as will the policies for storing and archiving data. But for now, Gadamer's theories allow me to analyze play analytic features from an historical perspective in order to determine how shifts in our prejudices towards game data alter how such systems may operate in the future.

Studying Play Analytics

Data analysis is both a serious and playful activity. It allows analyst to be normative in their pursuit while attempting to find insights and patterns useful to explain the dataset they are analyzing. Analyst can also take a non-normative approach to data analysis where analysis can be seen as existing for itself or allows an analyst to create their own methods of analysis. Play analytics, which applies data analysis to games, can

be viewed in a similar manner, both as a serious and playful domain. Play analytic systems provide players extra-diegetic experiences, set alongside, or completely removed from, real-time gameplay. Players can use these systems to optimize their gameplay strategy or simply use a system to reminisce about past gameplay experience. Play analytics affords both serious and playful data analysis.

In order to study play analytic systems both the design of the systems and the experience players have with these systems need to be discussed. As Bogost and Monfort layout in their platform studies framework (2009), there are different levels of analysis one can take when studying new media systems. Levels like the system's platform level, upon which the system is built, the form and function level detailing the design of the system or the reception level where users finally use the system, each express some of the levels one can take when conducting new media studies. In this dissertation I explore how to understand play analytics from both a design and experience perspective. This means I look at how developers design the 'form and function' level of a play analytic system compares to 'reception' level, meaning how players experience play analytic systems. To do so, I introduced the functional and phenomenological framework to explore the design and experience aspects of play analytics. For exploring the design of these systems, I use Bowker and Star's theories listing the properties of a categorization (i.e. data analysis) system: visibility, comparability and control. For exploring the player's experience of a play analytic system I use theories from phenomenology. Theories from three different phenomenological philosophers (Husserl, Heidegger and Gadamer) are used to approach describing how players experience play analytics as an act of playing with data.

Approaching the study of play analytics I ask two questions to uncover the properties of play analytic systems and the effects these systems have on a player's experience.

What is the state of the art of play analytics?

Currently, there has been no large endeavor to record and categorize examples of play analytic systems. Leaderboards, game databases, action statistics, player dossiers, recommendation systems, maps and replay viewers are all examples of play analytic systems. Given the variety of play analytic systems that exist it is the first task of this dissertation to examine these systems and the features they provide. Once we understand the breadth of what these systems offer only then can we ask further questions of how they affect gameplay and players. Knowledge of what play analytic systems are available not only help game developers understand the growing domain for these systems but also aid players who wish to produce their own play analytic systems.

Categorizing play analytic systems cannot stop at how these systems finally disseminate data. A thorough review of play analytics must include an exploration of how data is collected, how that data is parsed or otherwise made available, who has access to raw data and finally how a single data source may be used to produce a variety of play analytic systems. Only by studying the capabilities and limitations of these systems, and related system components like APIs, which are application programming interfaces that allow programs to access data, can play analytic systems be truly explored. This is why the first contribution of this dissertation is to provide an overarching analysis of play analytic systems. Defining what constitutes a play analytic system will enrich the other research questions in this dissertation which discuss how play analytic systems alter gameplay and how those same systems may be improved upon in the future.

How do play analytics alter gameplay?

Play analytic systems offer players the ability to analyze their past gameplay. Players can use these systems to reflect on their play style, indulge in hilarious moments captured during play or build their own systems capable of exploring their data. This alters how players approach gameplay. This is not to say that players have never before dissected gameplay outside of the actual game but with play analytics it is becoming easier and prevalent. How does increasing access to game data affect the player's experience? Does gameplay become merely a series of data points to count or does play

analytics offer different ways to approach playing games without actually being in a game? We must understand how play analytics alters games if we are to determine how these systems can benefit games and play.

In this dissertation I discuss how gameplay is altered due to play analytic systems which, I argue, work in a similar fashion to classification systems. Bowker and Star postulate that there are three properties of a classification system: comparability, visibility, and control (2000). Category systems functionally allow categories to be compared to each other, highlight (or make visible) specific defined actions to be categorized and enact a level of control over how categorized items are listed. I argue that play analytic systems also take on the same functional properties as categorization systems. Game data in these systems are compared, made visible (or validated through the use of game mechanics) and controlled. For example, when high scores are compared between players online those scores are made visible to all players and controlled using the rules of the game and the process to upload the scores online. Determining the effects of the functional properties of play analytic systems is therefore one way to answer the question of how gameplay is altered for players using those systems.

Another way to study how play analytics alter a player's gameplay experience is to take the same three functional properties of play analytic system but to approach them from a phenomenological perspective. If we view comparability, visibility and control as how a player experiences them we instead get the three properties of interconnected, exposed and historic. As I mentioned in the opening section of this introduction, systems like Steam exhibit these properties and in fact all play analytic systems do. Players do not see the underlying algorithms that compare their data together, until they compare their data with their friends. Players do not notice that their data is visible until they find out their data is exposed through a clever ruse. Players do not see how their data is controlled until they explore their gameplay history and find what is available. As I explore play analytic systems I argue:

Play analytics alter how players interface with a game's system and the identities of other players, increasing a player's ability to view gameplay as historic, interconnected and exposed.

Playing games such as Halo Wars or World of Warcraft allow players to interface with each other and the game's internal system. Play analytic systems offer new interfaces, ones that alter the perceptions players gain in-game. These interfaces provide the means of exploring historic, past gameplay and the interconnections between players create new perspectives different from perspectives gained while playing a game itself. This historic, interconnected data is exposed in both positive and negative ways, and forces us to review what play analytics is capable of beyond recording gameplay.

Overview of Dissertation Chapters

The following is a list of the chapters found in this dissertation complete with a summary paragraph describing each chapter's contents.

Chapter 2: Deconstructing Data

Beginning with data, this chapter covers how I define data and how it represents our environment. I critique the argument that defines data as a meaningless substance devoid of context or usefulness. Instead, data is presented as the initial stage of meaning for describing our environment. It is an initial stage along a spectrum connecting information, knowledge and understanding. Transitioning to each stage enriches data, through the data enrichment process, adding further meaning. I explain how game data relates to this enrichment process before providing an example of how my theoretical framework can apply to game data in the context of a play analytic system.

Chapter 3: Envisioning Visualization

Progressing from analyzing data, chapter three focuses on the domain of visualization. I argue that visualization, in its most basic form, is relating data using visual representations and therefore creates information (i.e. related data). There is a wide variety of visualization domains and ten of those domains are covered in order to give a

brief look into how visualization researchers approach visualization as a topic of research and a design practice. I relate these visualization domains to the data enrichment process and discuss in detail how the functional/phenomenological properties of my theoretical framework are used to interpret visualization systems (which includes play analytic systems).

Chapter 4: Analyzing Game Analytics

Chapter four goes beyond data collection and visualization to compare how game analytic and play analytic systems function as part of the game development and playing process. Examples of how game data is collected from digital games and used as part of the game development process are discussed. This includes looking at how analytics are used for business, design and programming purposes. From the player's perspective, I cover a number of play analytic systems built for players. Some tools are produced by game developers but others are player made, making some instances of play analytic systems completely player driven. Both game developers and players are discussed as having adequate need for better analytic tools, in general.

Chapter 5: Play Analytic Content Analysis

Chapter four covers the content analysis I performed covering current play analytic systems. Over seventy systems are analyzed to find commonalities and outliers in order to find what each type of system has to offer players. Some of these systems are small, mainly consisting of leaderboards, but others are quite large, such as systems that track millions of players like Battlelog or Bungie.net. Some systems provide ways of sharing user created content like Little Big World of the Sims Exchange, while others support exploration by marking game maps such as WoW Maps. What all of these systems have in common is that they exist as extra features in relation to the gameplay experience, provide players with additional means of becoming engaged with their games. Every system is discussed according to the functional and phenomenological properties that define both play analytic and categorization systems.

Chapter 6: Play Analytic User Study

Results from studying how players react to play analytics are delivered in chapter five. A system which tracked a number of game players as they played an online Flash game is used to review how players behaved during their play session. Players were asked to play a game for about 10 minutes before being shown one of three different play analytic interfaces (one which compared their data to another player, one which only showed the player their game data, and one which did not show any data) before being asked to play the game again. The differences between the three different study groups are reviewed along with their survey answers before and after taking part in the study.

Chapter 7: Conclusion

Ending the dissertation, I provide my final thoughts on play analytics, where I believe the domain is heading and what that means for the games industry. I discuss other examples like the Halo Wars story to question a number of problems and trends that exist within the domain. This includes describing play analytics as surveillance systems but also systems that allow players to critically approach gameplay.

CHAPTER 2

DECONSTRUCTING DATA

Defining data may seem like a rudimentary place to start exploring play analytics because society functions just fine without a firm definition of data. Data is the photo I took of my family last week, the expense report detailing this quarter's manufacturing costs, and the current temperature outside. Each of these examples hold claim to the term "data" and therefore we accept them as data. They represent description about our environment, numerical quantities or other attributes (Bergeron, 2002) and are symbols we can use (Ackoff 1989). We do not need to have an intricate understanding of data in order to use the objects we call data. Likewise, information is a term both given multiple meanings in our society and often used in concert with data. Information is "data with meaning" as some argue (Chen et al., 2009) whereas data solely represents raw facts devoid of meaning. Data is seen as being out of context and is only given context, and given meaning, through adequate processes such as relating data together (Ackoff, 1989). However, when we accept these overall definitions of data and information, issues arise that question what data actual represents and how idiosyncratic our interpretation of information actually is. For example, defining information as 'data with meaning' runs the risk of defining everything because we as humans constantly interpret the meaning of data and turn data into information by adding context to it. We cannot look at an object without making assumptions about the object, we cannot gather data without giving it context. Therefore, everything becomes information automatically, but what happens to the concept of data? In this chapter, I highlight these inconsistencies in the definitions of data and information, while discussing how these terms connect with one another. This helps me point out key concepts regarding how data and information are represented in analytic systems, such as those systems found in the domain of play analytic. Later in this dissertation, the concepts of data and information are useful to describe how data is represented in a play analytic system and how a system attempts to create information.

Defining Data

In the book Information Systems: Critical Perspectives, author Bernd Stahl applies critical theory approaches to information systems (2008). Stahl questions the interpretation of information as “data with meaning” and describes two problems caused by this definition.

“First, there is the problem that data is not simply brute facts of the world but that all data is already processed and gathered. Information thus cannot simply be the injection of meaning into data, because data already has meaning. ... Another problem of this definition [i.e. information is data with meaning] is that it renders information completely idiosyncratic. Data that may hold meaning for you may be utterly meaningless to me. This would contradict the implicit assumption that information is more generally accessible, which is required for it to be processed by machines.” (Stahl, 2008 p.71)

Stahl’s problems coalesce around data, information and meaning; “The difference between data and information is thus a difference in the level and appreciation of meaning” (Floridi, 1999; Stahl, 2008). In his first problem concerning the definition of information, Stahl argues that data inherently has meaning because it is collected and processed before any ‘extra meaning’ can be added to it. In essence, data is never neutral. Data, when referring to an entity, may be collected, processed and interpreted in a number of ways. Choosing a specific way of collecting, processing and interpreting the data gives it meaning. An animal for example can be seen as endangered, having a location, being a mammal, having aged to become an adult. Each of these data points carries with it a number of assumptions and requirements we must be aware of. We need to know what it means for an animal to be endangered, how to clarify location, define categories such as mammal and the animal’s lifespan. Even numeric data, such as temperature, is built on inherent assumptions. Knowledge of temperature scales like Fahrenheit and Celsius are required and knowing which scale is preferred within a culture is also helpful. Critiquing the definition of data and declaring data has meaning before

being turned into information allows for an analysis of how data provides meaning separate from information.

There are two ways data can be seen as having meaning. First, data is a representation of our environment. As Stahl mentioned, data has inherent meaning based on how it is collected and processed (2008). The methods for collecting and processing data determine how it is represented and I argue there are three properties that define the representation of data. The first property is physicality, meaning a representation of data has a form existing in the world (which can be both tangible and intangible entities). Words written on paper, the color of a car, a frame of video are all example of physical data. The physical attribute excludes cognitive concepts, like thoughts, and other concepts like processes or actions from being data (however, something like a diagram of a process is considered data). This is not to say non-physical, cognitive concepts have no bearing on data or their interpretation, just that they should be seen as separate concepts when compare to data (I discuss how thoughts affect our perception of data in chapter three for example). The second and third properties of data are awareness and referential. We must be able to become aware of data and be able to refer data to others. Awareness implies the ability to interpret an entity and understanding of how data is being represented, or at least an attempt to understand possible. A word, for example, is a physical representation of symbols (letters). We become aware of those symbols, acknowledge the sequence of symbols forming the word, and the ways of interpreting the sequence of symbols. Referring to data means acknowledge its existence as a description of our environment – whether we agree it is an accurate description is another matter – and understanding how to access it. We refer to data when we use it and can refer a piece of data to others, making them aware of the data. A dictionary refers a reader to the words that exist within a language and refers to the related meanings of those words. Becoming aware of data and being able to refer to it are also made possible because data is physical, an external entity to ourselves we become aware of and refer to. These three attributes work together to create a representation that has meaning. A representation gains meaning through perceiving the physical manifestation that we as humans become aware of and can refer to in our environment. This meaning may not be immediately

useful for completing a task, which is where the concept of information appears, but it is still meaningful to become aware of and refer to a physical piece of data.

The second way data is meaningful is how data is personally meaningful to us as individuals. A photo of my family holds more meaning for myself compared to a complete stranger in part due to the memories I have of my family. In order to define how data gains meaning in this way I return to the definition of meaningful I mentioned in chapter one. Meaningful data must “provide” something for someone. Data can provide utility or usefulness (i.e. normative) given a specific context. Knowledge about a company’s financial earnings can provide insight into whether the company can afford starting a new project. At other times data can provide something non-normative to someone, for example the ability to reminisce about my family in the photo. This is where non-data concepts like thoughts can provide further meaning to data. Our thoughts, ideas, and emotions can shape how a representation of data is meaningful to us, even though thoughts, ideas and emotions are not physical entities others can become aware of or refer to. And these meanings, shaped by our cognitive perceptions, may be deeply personal and not provide the same meaning for someone else. As Stahl argues, “data that may hold meaning for you may be utterly meaningless to me” (2008).

As Stahl notes, the fact that the meaning of data can be altered by our individual perceptions of data makes the meaning of data incredibly idiosyncratic. What data provides me may not provide the same thing for someone else, and may provide something entirely different. My medical records provide me with acknowledgement of my health while the same records may provide my doctor with an understanding about how my health may progress over time and into the future. Everyone can approach the same dataset but find different meanings within that dataset. This phenomena is encapsulated in the theories defining the term boundary object (Star and Griesemer, 1989). An object is a boundary object when one group of people finds the object meaningful for one reason and another group finds the same object meaningful for different reason. The representation of the object is the same but the meaning changes between groups. A boundary object often allows groups to communicate with each other,

since the object can act as a translation point, even though they have different ‘meaningful’ associations with the object. The idiosyncratic nature of interpreting a boundary objects can thus be a useful tool, aiding in the sharing of perspectives, interpretations and meanings. Boundary objects provide a method of communication to other groups on top of object providing value based on how the object is interpreted in each particular group. If something like a medical record, which can be considered data, can function as a boundary object then other forms of data can act as boundary objects too. Therefore, the idiosyncratic nature of data should not be seen as a hindrance but as way for multiple meaningful interpretations to be shared about a piece of data.

Data relies on both representation and meaningfulness (which implies the idiosyncratic notion of ‘providing for’) to gain meaning. I combine these two properties to form the definition of data I use in this dissertation: *data is a representation that is meaningful*. How data represents our environment and what data provides us form the concept of data. Now that data is defined I can return to discussing the connection between data and information. Separating data from information allows data to be seen as having meaning and information as a concept that alters meaning. We can become aware of data, refer to data and prescribe meaning to data but turning data into information means taking further steps to form relationships between data.

Data Transition and Enrichment

One framework used to describe the connection between data and information is the data-information-knowledge-wisdom (DIKW) hierarchy (Ackoff 1989, Chen et al. 2009, Shedroff, 2001) The hierarchy is used by designers and researchers to discuss how one *phase* (i.e. data, information, knowledge or wisdom) of the scale moves, or *transitions*, to the next *phase*. For instance, data moving to information is a transition between those two phases. The general consensus is when a phase transitions to the next phase meaning has been added to the previous phase contents (Ackoff 1989, Chen et al. 2009, Shedroff, 2001). Data is given further meaning when it can be used to answer hypotheses and becomes information (Ackoff, 1989). Information is given further meaning when it can be combined with other perspectives and becomes knowledge

(Shedroff, 2001). These are just a few examples but every interpretation of a DIKW hierarchy is concerned with defining the phases and transitions between the hierarchy elements.

While my definition of data is not defined in the same way as data is defined in other DIKW hierarchies, the hierarchy still applies. Even though I argue data has meaning, adding meaning to data still transitions data along the hierarchy from information to wisdom. The only real difference is how the transition between data and information phases work. I define information as a *relation between data*. This is perhaps the most basic and broad definition I can use to refer to information. Data can be given a different context through relating it to other data, thus adding further meaning to data and creating information. A relation informs someone of how data can be correlated or combined with other data. When I learn the amount of horsepower a particular vehicle has I am learning the relation between the vehicle (one piece of data) and the horsepower attribute of that vehicle (another piece of data). It's the relation which provides further meaning to the data. This definition of information conforms to other definitions that describe information as the ability to answer questions (Ackoff, 1989), the results of a process (Chen et al. 2009) or the process of providing context (Shedroff 2001). The only major difference between my definition of information and the other example definitions is the assumption that data initially has meaning. Therefore, when I speak about turning data into information I can talk specifically about what transitions occur to add further meaning to data that already begins with meaning.

The other phases in the DIKW hierarchy are knowledge and wisdom. I define knowledge as a *complex configuration of information*. This definition is the same as Shedroff's definition of knowledge (2001). By relating data (which creates information), and combining multiple relations from different contexts, knowledge is formed. For example, once I know more features about a vehicle's engine and how the car performs I gain knowledge about that vehicle. For the Wisdom phase I use the term understanding instead (as some DIKW hierarchies do). I define understanding as the use of knowledge in unforeseen or novel situations. Similar to Shedroff's definition of wisdom (2001),

understanding represents when knowledge is taken farther into situations someone did not foresee. Taking my knowledge of a car's engine can turn into understanding if I successfully build my own engine. Even more so, building an engine configuration I have never seen before would constitute an understanding of engine manufacturing.

While I mainly rely on the link between data and information in this dissertation, I do acknowledge the transitions to knowledge and understanding. Transitioning to knowledge means someone has seen multiple data and information combinations from different perspectives. If we return to an example from play analytics, a replay file (which details the events of a previous gameplay session) can be presented from many different perspectives using a replay analysis tool (Belicza, 2010). Information about how each player managed resources, which strategies players used and how players reacted to each other over time combine and give the player reviewing the replay a chance to gain knowledge of the replay events. If the player then applies that knowledge in a game they gained an understanding of the game through analyzing the replay. From a gameplay performance perspective, this is the ultimate goal for a play analytic tool; players become better at playing the game because analyzing past gameplay gives them a greater understanding of how the game works.

The DIKW, or in my case the DIKU, hierarchy is designed to cover how meaning is added to the hierarchy phases. What about the cases where meaning is taken away? My definition of data, for example, assumes data gains meaning from its representation and from what it provides someone (i.e. how data is seen as meaningful). However, as Stahl says "data that may hold meaning for you may be utterly meaningless to me" (2008). If a representation of data is incoherent or unimportant to someone then it is meaningless. This can occur if data is represented as a code or is hashed using an algorithm, for instance. The data still exists but is hard to read or impossible to read without knowledge of a decoding method. Data may also be meaningless based on my interests. I am not a car mechanic and have no inclination to review diagnostic reports detailing how a car operates. A car mechanic may find those reports useful but for myself, as an outsider, the data does not 'provide' me with anything. Additionally, a transition between phases –

between data and information, for example – can also decrease meaning. If a graph used to combine data is hard to read or the variables are labeled wrong the graph can decrease meaning. If a correlation is made between data but the viewer does not agree with the reasons for the correlation meaning can decrease. Throughout this dissertation I refer to the situation when data becomes meaningless or a transition causes a decrease in meaning as inan. I use the term inan to contrast the term for adding meaning, enrichment.

Enrichment and inan are used in this dissertation to cover how meaning is added or taking away from data, information, knowledge, etc. When I cover play analytic systems I discuss how specific features within those systems alter the meaning of data for players, by either enriching the data (adding meaning) or making it inan (removing meaning). Furthermore, I can use the concept of transition, the act of moving from one phase to another along the DIKW hierarchy, to cover when enrichment causes one phase to move to the next or if a transition causes an inan situation. Enrichment and inan can also be used in combination with the functional and phenomenological framework. The functional property comparability, for example, describe the methods and processes for enriching data in an analysis system. Phenomenological properties like historic describe how data may become inan over time if a user only views a dataset as meaningful in present circumstances. How a system functions and how a player experiences a play analytic systems both have an effect on how data is enriched or made inan. But before I provide an analysis of how the functional and phenomenological framework can interpret game data, based on the arguments I've stated thus far, I present examples of datasets found in games. These examples give an initial overview of the type of datasets play analytic systems use and what data players actually experience.

Game Data

While the DIKW hierarchy can apply to any type of data, in this dissertation I specifically talk about *game-related data*. Any data related to the game development process or playing a game is considered game-related data and this covers a wide array of possible datasets. For example, the following game industry professions and players can use data related to the following areas:

- Community Manager – Social connections between players inside a game environment and other online community areas.
- Financial Manager – Game and micro-transaction sales.
- Game Designer – Recorded player gameplay associated with a game’s mechanics and interactions among players.
- Producer – Production schedules.
- Programmer – Source code complexity and hardware performance.
- Player – Individual gameplay, game content and “exquisite gameplay” (Medler, 2009).

In this dissertation I focus on the game-related datasets that players typically come into contact with while using a play analytic system. Datasets related to the game development process or the hardware performance of game servers are not the typical datasets found in play analytic systems. Instead, players are interested in datasets that relate to playing a game. A player’s gameplay data – their scores, achievements, statistics – is one example of the type of data related to playing a game players may wish to use and analyze. A game’s content – models, levels, quests, characters – is another example; players may wish to interpret what content is available in a game because the content is what the player actually experiences. Additionally, play analytics is about analyzing data away from gameplay. Play analytic datasets are available to players outside of the game’s environment. The data I discuss is therefore typically recorded or created to represent one of these two areas: a game’s content files (and any additional supplemental content created by players) or a player’s past gameplay.

The types of game-related data often found in play analytic systems include (but are not limited to):

Assets

Game content such as image sprites, 3D models, levels, textures etc. Any content used within a game representing “physical” objects, in the context of a game environment, are assets.

Components

“Non-physical” game content related to playing the game, and cannot be referred to as an asset. Examples include item statistics, text describing quests, win conditions or goals, and other data covering the rules of the game.

Gameplay Events

Happenings or occurrences related to a player’s, or group of player’s, gameplay. Events can be temporal based, such as completing a level, or can accumulate over time, such as the number of levels completed.

Gameplay Statistics

Values related to how the player is performing in the game. For instance, scores, ratings, and rankings on a leaderboard are all considered types of performance interdictors. Statistics is also a commonly used term to refer to all types of player data.

Maps

Data used to describe virtual space found in games. Maps are often available as part of a game’s environment but are often used outside of the game to provide spatial context to game-related data.

Replays

Records of past gameplay session used to recreate the gameplay. A replay allows viewers to “re-experience” the gameplay recorded. Replays may work asynchronously like a linear recording of the game, where viewers fast forward or rewind time, or replays can be replayed synchronously during another player’s gameplay (these are often referred to as “player ghosts”).

Save Games

A recording of a game's current state that is used to initialize a game to mimic the game's configuration when the save was captured. Data related to the player's avatar, past game event history, changes to the game environment or other aspects that alter how the game operates are typically recorded in a save game.

Stories

Narratives created by players related to their gameplay. These narratives or stories do not need to be directly related to a game's storyline, as some games allow players to create their own stories or histories for their characters.

Videos, Screenshots and other Media

Video, images or other media captured during gameplay. Some games automatically capture media like images or video during gameplay and other times the player specifies when the capture occurs.

The two major factors defining game-related data, found in play analytic system, are whether the data is captured from a player's gameplay or the data represents game content. Player gameplay includes gameplay events, gameplay performance, replays and saved games (referencing the list above). Game-related content includes models, components, maps, stories and strategies. The method I use to categorize these game-related datasets is based on two criteria: a temporal relationship is used to categorize gameplay data, and how data is produced categorizes game content.

Time is a major factor when it comes to recording gameplay data because a) games are played over time and b) different time frames (single moments or long stretches of time) can be represented as data. Player gameplay datasets can fit into three categories based on their time frame: concurrent, progressive, and longitudinal. Concurrent data collection happens at the moment an event occurs in a game. An event is represented by any action or decision that is made by the player or the game itself. For example, when the player dies (most likely due to a failed decision) a game records that

event as a loss of one life. The recording of that event is a concurrent recording type because it happens when the event occurs. Other examples include: earning an achievement which represents a onetime event, capturing a screenshot, or winning a level. Concurrent data typically contains a time stamp to denote when the event occurred in time.

Progressive recording is a string of concurrent recordings. Each time a player dies a concurrent record is made and the total amount of player lives over time is the progressive record. Another example is a player's score. While each point a player acquires is a concurrent record, the final, total score is the important record representing a progressive recording across time. Also, progressive data may literally record a player's progression over time. Replay files contain a record of every occurring movement and action in a game and players are given the option to review, rewind and fast-forward through their gameplay. Gameplay videos are also progressive data and act similar to replays. Progressive data can either be seen as an aggregation over time or a literal portrayal of data over time.

Longitudinal data represents data a game can use between game sessions and still has value to review outside of the game experience. A saved game file is longitudinal data. The file can be read in by the game at a later date and the files contents, the data contain within the file, can be analyzed outside of the game. Game content is another example. Models, maps, items, and quests can each be recorded for outside analysis but is also data a game must use to create the gameplay experience. Concurrent and progressive data, when discussing data collected for play analytic systems, represent data that is not reused by a game. A replay of a game, for example, is not replayed while a player plays the game, they must view the replay outside of gameplay. Longitudinal data can be seen as having qualities of concurrent or progressive data (a game save is an aggregation of a player's experience over time) but longitudinal data can be reused by a game.

Beyond the three different time frames of data collection during gameplay there is also the data category known as external data, data created outside the gameplay

experience. Online game databases are filled with external information: walkthroughs, cheat codes, strategy guides, etc. These databases have taken the place of printed strategy guides, telephone hint lines and personal knowledge passed from player to player (Consalvo, 2007). Other examples include players writing their own stories about their gameplay or players recording data related to emergent behavior, such as creating a system for managing the activities of guild members (dragon kill points (DKP) systems are used by guilds to give guild members currency for participating in guild events but these systems are not tracked by the game itself). External data is accessory data to the gameplay experience but still relates to the game experience in some way, providing value to players.

Play analytics uses each of the four categories of data. The temporal data categories (concurrent, progressive and longitudinal) generally are used to gather data regarding a player's gameplay. Gameplay events, replays and saved games are all temporal based and are related to a player's gameplay. External data refers to the data related to a game's content. Models, components, maps, stories, walkthroughs and strategies do not necessarily have a temporal component but still represent game content players may find valuable.

Functional Data and Data Phenomenology

Returning to the functional and phenomenological framework I described in chapter one, I can begin using it as an interpretation tool to discuss my definition of data, the enrichment process and game-related data. For example, if we start with the functional properties, the visible functional property relates to the definition of data. Visible data is physical, something players can refer to and become aware of. Scores, resources gathered and units created are all instances of visible Halo Wars data, for example. The stat tracking system makes players aware of those values and players can refer to the values. Comparability, the second functional property, relates to the data enrichment process. Data is compared in order to create information and provide meaning to players. A player's Trueskill score (Herbrich and Graepel, 2006) in Halo Wars is created through a comparing process. Wins and losses are compared between players and

used to calculate a player's overall skill level, their Trueskill, according to the system. The third functional category, control, relates to how the developers/maintainers of a play analytic system controls the visible data and how it is compared. 343 Industries, the developer of Halo Wars, controls the Halo Wars stat tracking features. They decide whether a player's score is shown, how it is presented and how it is compared. Players have little say over what data is collected, presented and used in the system. While the developer may not have any overt malicious intent in regards to having complete control over the stat tracking system, players have a greater difficulty forging their own meaning from the data beyond the developer's control.

From a phenomenological perspective, data experienced in the context of Halo Wars, for instance, revolves around the competitive nature of the Halo Wars community. Competition begins with a clear distinction of who is beating whom. In Halo Wars, a player's highest Trueskill score and match scores are the competitive distinction. These score values need to be exposed (the first phenomenological property) in order to facilitate competition or at least this is what the Halo Wars community has decided. They fought to keep the stat tracking system up and running, players wanted their score values exposed. These players had a good reason too. Their scores and win/lose records are the means by which Halo Wars players interconnect (the second phenomenological property). Competition is their connection through gameplay. They of course have forums and other communication outlets where they interconnected but these are not associated with actual gameplay. Players have continued to play Halo Wars for the last three years because they have had the means to compete with each other as a community. The Halo Wars stat tracking system represents an acknowledgement of the competitive community, allowing players to interpret their data from a competitive mindset (we can say they have a competitive mood). Additionally, the stat tracking system provides players with a history (the third phenomenological property) giving them the ability to look back at their past accomplishments, how well they performed over time and how their individual matches unfolded. This history creates an identity for each Halo Wars player. The players at the top of the leaderboard are seen as both skilled players and opponents that need to be conquered. As time passes, new players rise through the ranks

and the historic data defining each player grows richer. Analyzing a play analytic system using the three phenomenological properties allows us to interpret, and question, why players use these systems and how the functional properties (like the decision to visualize scores) affect how players interact with the system.

One can also critique the Halo Wars system using the same functional and phenomenological framework. Player histories, for instance, are only really made valuable through a player's rank on the leaderboard. Players have few ways to gauge their progress over time besides viewing individual matches in succession. Adding some sort of timeline feature or method of comparing a player's progress over time may help players further evolve their skill sets. A timeline system could be built to show how effective a player is when using specific build orders or strategies. If the system made that type of information visible to the player they would be able to determine which strategies they need to practice more. Players also have a hard time analyzing and interpreting the data from a mindset other than a competitive one. As I continue to discuss play analytic section in future chapters, we see examples of systems that try to provide players with multiple avenues for interpreting their data (not just from a competitive perspective) and how the functional and phenomenological framework can apply to these play analytic systems as well.

Summary

What I have tried to show in this chapter is that data is a representation that is meaningful. Data is never neutral and is always collected with a purpose; where that purpose is typically meaningful to someone. However, this purpose is usually only meaningful to the one who initially collected the data or set up the system to collect the data. In order to make data more meaningful to other people, to turn data into an object that 'provides' something for someone, we need a system for enriching data, to add meaning to the data. The DIKU hierarchy, which is used by many information/visualization researchers, is one such way researchers describe the process of altering the meaning of data. Enriching data turns data into information, by comparing data together, and enriching information creates knowledge, by viewing information

within different contexts. Although, sometimes enriching data can lead to data becoming inan, or meaningless. Leaderboards, for example, can enrich data for some (by comparing their score to other players' score and create a sense of competition) but may make data more inan to other players (for those who do not care about their score being compared to other players). In the end, we must realize that data is idiosyncratic and can provide something different to each person. This means play analytics, a domain dealing with game-related data, needs to understand the idiosyncrasies of players and determine how data is made meaningful to them.

CHAPTER 3

ENVISIONING VISUALIZATION

Representation and meaning, two interlocking parts of data, the combination of the physical representation of our environment and the meaning of those representations manifest themselves within data. As we saw in chapter two, there are plenty of ways data is represented and provides meaning in games. Game maps are visual representations of a game world and provide an overview of places a player can explore. Values depicting score are both numeric representations and a method to antagonize players to perform well. Written game lore are textually represented allowing players to explore extra content describing a game's world and characters. While there are other data representation possibilities – audio and tangible for instance – vision is the sense that we use most often to experience the representations of maps, scores or text. The ability to see and acknowledge the presence of data, is a powerful tool for creating representation and making data meaningful.

Adding meaning to data by using visual representation is where we proceed next. While other senses such as hearing and touch have a strong case for being vital to gameplay, vision is almost universally used in every form of game, digital or otherwise. Our brains are hardwired for visual processing more so than any other sense (Galotti, 2010) and visual representation can be used in a variety of ways, a variety that affects the meaning of those visuals. This chapter covers how visual representation is used to relate data in order to create information and how visual representation is used to counteract data becoming inan (i.e. meaningless) when users are faced with increasing amount of data.

Play analytic systems share a common problem with other data analysis systems; there is typically a large amount of data available to players and the developers of play analytic systems need to create methods that allow players to find meaning amidst that data (i.e. methods for enriching data). I argue that the process of *visualization*, which is

defined as relating data through visual representation, provides the methods for enriching data, as I discussed in chapter 2. Visualization can provide the mechanism to create information within a play analytic system by visually relating data. In order to explore how visualization is used I examine a number of visualization methods covering how they represent and relate data. These visualization methods can be understood in both a functional and phenomenological way, where visual representation affects the context of data being presented to users and how those users interpret (or frame) the data being represented. Understanding visualization through a functional and phenomenological lens allows us to connect data and visualization together, which are two necessary components for defining play analytics.

Your Vision

Vision is one of the most important senses we as humans have for acknowledging and interpreting the environment around us. Our brains have a large “neural investment in processing visual information” (Anderson, 2004) with many cortical and subcortical regions of the brain devoted to processing visual data. Stimuli from our environment, often called distal stimulus, is projected through our eyes as proximal stimulus, a replicated image of what we see in the environment (Galotti, 2010). The cortical and subcortical regions in our head are used to interpret the proximal stimulus, the replicated image of our environment.

Percepts, as they are called, is the term used to describe mental interpretations of the stimulus we receive, the proximal stimulus, interpretations which do not necessarily have to conform to the literal stimulus being visually sensed. For instance, when moving slightly towards and away from an object the actual size of the object, the distal stimulus forming the proximal stimulus we sense, changes based on the distance between the object and ourselves. An object far away is smaller and bigger when close. Try stretching your arm out and moving your hand slowly towards you. The size of your hand should appear to stay constant, or so subtle that our brain dismisses the change in size (a phenomena called size constancy (Galotti, 2006)). This is an example of distal stimulus creating an optical illusion where the stimulus we see represents one thing but our brain

interprets the proximal stimulus (the projection) as another. Then there are cases where the distal stimulus creates no illusion but memories relating to the stimulus affect the mental associations we have while experiencing the stimulus. When you look at a photo of a friend what are your thoughts as you gaze at the photo? If you thought about the proximal stimulus being projected through your eyes you may notice the color intensity of the photo or the outline of your friend against the background, not the last time you saw your friend or the event when the picture was taken. We can think about the former but most of the time we think of the latter, not noticing the physical aspects of the stimulus and instead focusing on other mental associations with the visual image (the percepts). Focusing on mental association means our perception of our environment is not always of the literal environment stimulus but includes any number of other factors related to the visuals we are receiving. Both the proximal stimulus (which is a reflection of the distal stimulus from our environment) and the user's perception of the visual data are factors in presenting and understanding the visual stimulus, and visual data.

Visual representation can take advantage of visual perception concepts such as those found in Gestalt Theory (Galotti, 2006). Relying on concepts like proximity, similarity, or continuity, Gestalt theory describes how visuals can be used to relate data. Standard Cartesian coordinate charts, for instance, contain two axes (one horizontal, one vertical) for mapping two data points, or variables, onto each of the axes in the coordinate systems (Figure 3.1). Cartesian charts visualize data by relating two datasets together through proximity, organizing the datasets in a way that spatially connects them. Once organized in proximity to each other, various other ways of relating data can occur. Using color, size, position, shapes, symbols, or labels, additional variables and other attributes about the datasets can be related as well. In fact, it is the act of relating data together which I use to describe the process of visualization. Visualization means relating data through visual representation. However, not all visual representations are visualizations.

Visualization must, like visual representation, be physically present, and, unlike visual representation, the relations between data must be physically present as well. A representation, as defined in chapter 2, must be physical, and we must be capable of

referring to and becoming aware of the representation. Physicality describes an entity with an existing form in the world (this excludes cognitive concepts like memories). We can become aware of these entities and refer to them as outside ourselves. However, a visual representation does not have to relate data. This is how we can differentiate visualization from visual representation alone. A picture of a dog is only a visual representation. It physically exists; I can become aware of and can refer to the picture. The picture is not a visualization because there is no physical relationship presented. I may notice that the dog in the picture is similar to my dog and infer other data points such as the dog's breed but these are only mental, non-physical relations. I cannot refer others to my mental relations I am creating with this picture nor can others become aware of the relations. Relating memories to the picture does describe a way of relating data, of creating information, but it does not represent a physical relation others can use. Thus, when I am talking about visualizations I am talking about a representation of an entity that includes a relation others can become aware of and refer to. Visualization is a way of communicating relations through external entities.

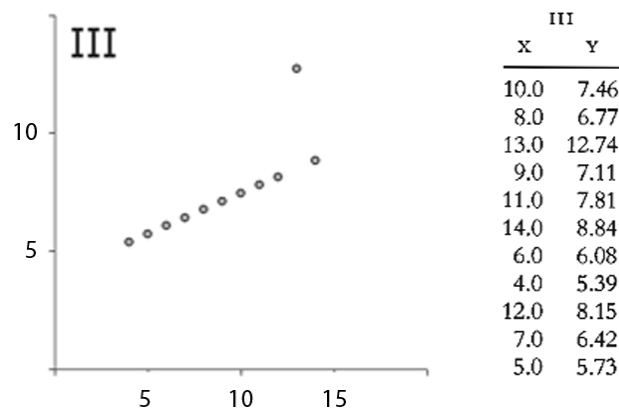


Figure 3.1: Mapping data points onto a graph can express patterns which would otherwise be harder to see. (Graphs reprinted from (Tufte, 1983))

Visualization, as a practice of relating data together, is often proposed as a normative, utilitarian method of creating information. Borrowing an example used by Tufte (1983), Figure 3.1 displays both a table of variable pairs and a Cartesian chart with the same variables mapped to the two axes of the chart. The chart makes it easy to notice visually that one of the variables is different compared to the others. Most of the variables follow a single line but a single variable jumps out and does not conform to the curve.

Looking at the table of values it may be difficult to immediately see which variable deviates from the pattern set by the other values but the table does provide the benefit of allowing me to find the exact numeric values of the variable outlier once I recognize it on the chart. Both the table and chart are representations of the same data, attempting to portray an objective view of the data, and both have their strength and weakness when it comes to determining patterns, similarities or other relations.

Our experience with data relationships presented through visualization does not solely rely on the normative aspects of visualization, however. Relating data may be personal, may be ambiguous, may be uncertain. Color for example is cultural and can affect how someone interprets a visualization. Where red symbolizes good luck in Asian cultures, red is typically used to mean danger or debt in the United States. If I were to build a system that visualized a company's lost revenue over the year, red is a proper color to use for an American company but perhaps would not be the best color choice for an Asian company (Ware, 2000; Hall, 2008). Our own interpretations of data and other factors such as cultural expectations can impact how we visually experience data.

Visualizing data means tapping into neurological aspects of our brains that provide meaning to data in addition to the physical aspects of the visuals in our environment. We can interpret physical stimulus as something existing in our environment, making use of factors such as proximity or continuity, and we mentally interpret stimulus by relate it to such things as emotions or memories. But these aspects can be thrown into chaos if we receive too much visual data for us to handle. If the chart or table presented (Figure 3.1) contained millions of variables our ability to comprehend the proximity or continuity of the data would quickly diminish. Likewise, if there is too much data to comprehend we cannot adequately connect it with other meaningful interpretations we may have stored in our memory as past experiences. Issues regarding proper data representation and how to give meaning to data become more prominent as we begin to deal with larger amounts of data.

Information Overload and Filter Failure

Acquiring data comes at a price. Concepts like “information overload” (Edmunds and Morris, 2000) are said to hinder decision making and decrease efficiency due to “receiving too much information” (Eppler and Mengis, 2004, p326), where information is often used synonymously with data. A few of the causes attributed to information overload in Eppler’s and Mengis’ (2004) review of the related literature include uncertainty of information – the level of ambiguity, novelty, complexity, and intensity (Schneider, 1987) – complexity of tasks using the information (Tushman and Nadler, 1975), and technology advancement such as storage capacity or ease of access (Schultze and Vandenbosch, 1998).

Games that allow players to create their own content can fall victim to information overload. Little Big Planet, Spore and The Sims each have large online databases filled with content created by players (levels, items, characters, etc.). There may be thousands, even millions, of created objects to browse, making it difficult for players to find meaningful content. Having too much data, even data that is deliberately collected, can be a burden to the user, causing information overload. Even though these databases tend to use a form of quality control, such as rating systems or tagging, there still tends to be too much information for any one player to view in order to find the content relevant to them. Perhaps the inability to adequately judge and filter data can be argued as the real cause of information overload.

Shirky argues that the problem of having too much data is not information overload but filter failure (2008), referring specifically to data we receive from media. In the past media outlets – publishers, broadcasters, filmmakers – filtered media for quality. With the rise of digital media and participatory culture those filters have ceased to provide adequate quality filters. The volume of data has increased but the ability to filter the data has not. If we had the necessary tools to organize our data the volume of data would not be as much of a hindrance. As Eppler’s and Mengis’ write:

“Simpson and Prusak (1995) argue that modifying the quality of information can have great effects on the likelihood of information overload. Improving the quality (e.g., conciseness, consistency, comprehensibility, etc.) of information can improve the information-processing capacity of the individual, as he or she is able to use high-quality information more quickly and better than ill-structured, unclear information.”(2004, p331)

The inability to process large quantities of information is not caused by the abundance of data but the quality of how that data is presented and accessed.

The “quality of data” relates to the concept of inan data I brought up in chapter two. Meaningless data, inan, loses the ability to provide something for the user or may have an inadequate representation. Acquiring too much data can cause users to become less aware of data and it may become difficult to even realize data exists. Even when a user is aware of data they may have no use for the data or have no experience or connection with the data, making it meaningless. A picture of a dog may be meaningless to me when I do not know who the dog belongs to or what breed the dog is or who took the picture. Viewing the picture of the dog means I accept the picture exists, I am aware of the visual representation of the dog, but beyond that fact the picture is inan to me, easily dismissed. Now think of this one instance of an inan picture on a massive scale. Millions of pictures are added to online photo sharing services like Flickr every day all of which are most likely inan to each of us. Granted, someone may find meaning in the act of looking at random pictures others have taken, thus not every picture is immediately inan. Perhaps random photos of beaches are meaningful to someone who grew up near a beach. How would one find such images though? The sheer number of photos that exist and continually added to a photo sharing service makes it inconceivable to hunt down personally meaningful photos through a brute force approach of viewing every photo. The problem is therefore a filter problem; if I wish to find photos that are meaningful to me then I need tools to help me find those photos in a large dataset. What methods can I use to enrich data, counteracting data becoming inan, and move towards information?



Figure 3.2: Mini-maps are used in some games to provide an overview of a game space and color code various objects within the space such as units or buildings (Starcraft 2's mini-map displayed).

Information, as defined in chapter two, is created when data is related to other data. Finding a connection between data creates a relationship, a new frame of reference we can use to interpret the data. A new relationship or interpretation can help increase the quality of data by removing uncertainty and be used as criteria for filtering out meaningless data. One example from real time strategy (RTS) games is the interface concept known as a mini-map. RTS games are often played in an environment that is larger than the viewable area players can see at any one time. Their screen acts as a camera in the environment but is only limited to a small area of view. A mini-map is a way to provide a player with data about the entire environment at once in as little space as possible. Figure 3.2 is an example of what a mini-map looks like during a match in the RTS game Starcraft 2. The small dots on the mini-map represent unit locations in the environment. RTS games often give players control over many individual units (soldiers,

vehicles, etc.), which are used to achieve objectives such as destroying other opposing units. Mini-maps color code a player's units making it easier to see where their own units are and where other player units are located. If units were not related by color on the mini-map it would be difficult when glancing at the map to know who owned which units, causing the uncertainty of the data to rise. This is not to say a mini-map without color coding does not relate data (the unit positions are still related to each other in a single area) but color coding is a better relation technique for increasing the quality of data since there is a useful aspect of the environment to map onto colors for the observer. A player glancing at the mini-map can filter out which units are friendly and can focus on finding enemy units, for example. Therefore, if I argue that play analytics is a domain which presents game data to players it is beneficial to understand the different ways game data can be related and filtered, particularly through visual representation. In order to do this, as I described the process of relating data visually in the last section, I examine how different visualization domains approach visually representing data.

Visualization Domains

Visualization uses visual representation to enrich (add meaning to) data through emphasizing relationships, which can increase the quality of data, and creates information in the process. This definition is one of the most basic definitions I have settled upon in order to explore visualization as a broad concept. I need the definition to be broad because data interpretation is a broad concept too.

I've already argued that both normative and non-normative approaches can be applied to interpreting data. If visualization is meant to add meaning to data, to allow further interpretation of data, then I need to be able to discuss a variety of approaches to visualization, both normative and non-normative. In order to discuss a variety of approaches to visualization, I present the related work from ten different visualization domains. These ten domains are not meant to represent the entire practice of visualization but they do provide a variety of methods to analyze (Figure 3.3). Some are the higher level visualization domains that are typically considered the main visualization disciplines: information visualization, scientific visualization and artistic visualization.

Other domains – casual, critical, narrative, personal and playful visualization – are sub-domains below the other main visualization disciplines (information, scientific and artistic) and visualizations built within those domains can be considered examples of visualizations from those disciplines (for example, a narrative visualization may be considered an information visualization as well). Finally, two of the domains, minimalist visualization and directed visualization, can be considered design philosophies and promote the methods for visually representing data. However, for the purposes of this chapter I present each of the domains as their own entity and let me describe how visualization allows for different data interpretations. After defining each of the ten domains I return to my functional-phenomenological interpretive framework in the next section and highlight how the functional and phenomenological properties apply to the visualization domains.

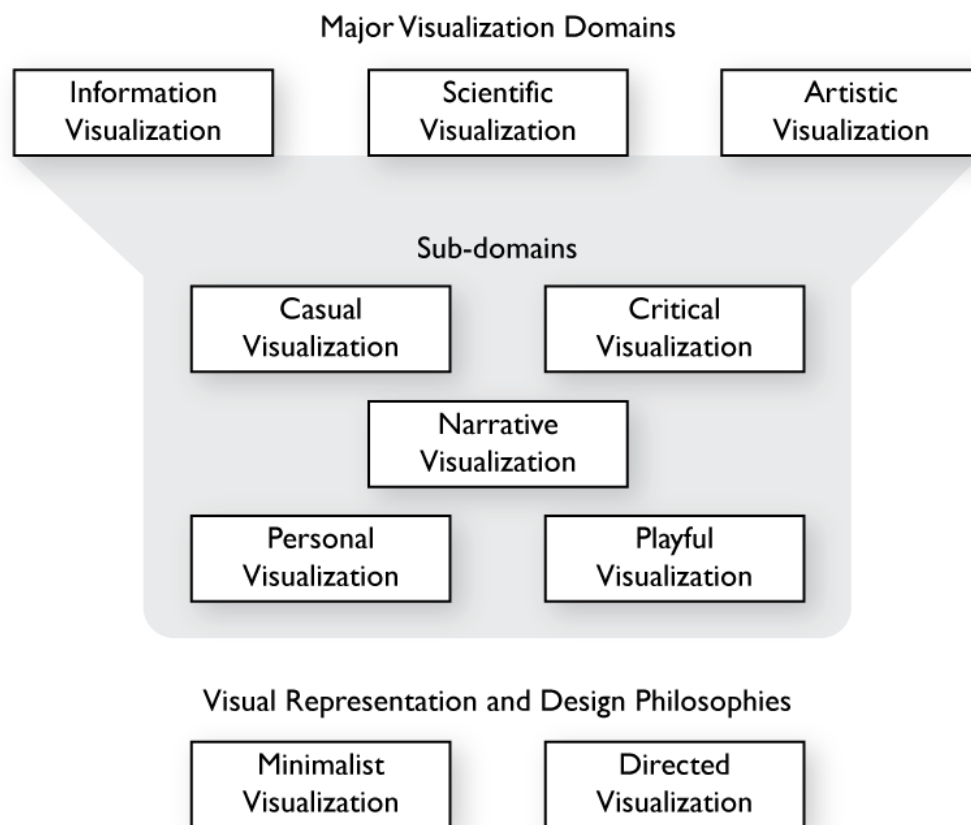


Figure 3.3: While I present ten visualization domains below, the domains are not exactly equal to one another. This hierarchy shows how some of the visualization domains are the major visualization

domains (or disciplines), some domains are actually sub-domains of the major visualization domains and a few are visual representation and design philosophies.

Information Visualization

Information Visualization (infovis) is the most referred to and discussed form of visualization. While data visualization can be argued as being the main overarching domain (Few, 2009), one which all other forms of visualization reside under, infovis is generally considered the default domain when referring to visualization. I should also point out that my definition of visualization is similar to ‘data visualization’ which has been defined as “visual representation that supports the exploration, examination and communication of data” (Few, 2009) but I argue that the processes ‘exploration, examination and communication’ refer to an acknowledgment of the relations existing between data.

While the definition of infovis can vary (Spence, 2000; Friendly and Denis, 2009) many have submitted to the definition presented by Card et al. in *Readings in Information Visualization: Using Visualization to Think* which is: “the use of computers to interactively amplify cognition using visual representation” (Card et al., 1999). Let us step through the definition terms: computer, interactivity, amplify cognition and visual representation. Infovis is seen as a computational-based domain, relating to digital technology, in contrast to other visualization domains such as information graphics, which are either made for print and, if presented digitally, are non-interactive. The ability to interact with a visualization is important because interactivity means a user has some level of control over how they experience the visualized data. If the purpose of visualization is to relate data, a user must be able to explore those relations. A user cannot rely on the fact that someone else will create a non-interactive (i.e. static) graphic that describes the exact relationship(s) the user is trying to examine. However, the terms amplifying cognition and visual representation found in the definition of infovis apply for both infovis and information graphics (and as we will see amplifying cognition is a major departure point for other visualization domains).

Few defines amplifying cognition as the “ability to think about information by assisting memory and representing data so we can comprehend it” (Few, 2009). In other words, infovis helps us think about data using visual representations, increasing our cognitive activity (Spence 2000, Ware 2000). We learn to comprehend the represented data and, hopefully, gain insight from the relationships being presented. The type of data usually represented using infovis is abstract, non-physical data. Non-physical, in this context, refers to non-tangible properties meaning infovis typically represents data without an initial visual form (Spence, 2000). Visualizing a skeleton to show the location of a broken bone or the floor plan of a building are considered tangible data, which typically fall under the jurisdiction of scientific visualization (which is presented next). Instead, non-tangible, abstract data can consist of quantitative data (food prices), events (stopping or starting a video) or attributes/categories regarding an entity (color, genre, length of time) that do not have a tangible visual representation.



Figure 3.4: Google Analytics (pictured) collects user behavior collected from websites and maps the abstract data to visual representations such as line graphs.

Visualizing abstract data means information visualization has to map data onto a visual representation (Manovich, 2010). Google Analytics is an example of an infovis tool where abstract user data captured from websites is mapped onto visual representations to show quantity or frequency values (Figure 3.3). The way the tool works is a developer who operates a website places a snippet of tracking code into each of their webpages in order to gather information about how users are behaving on the developer’s website. The code captures data from each user regarding their IP address,

the browser they are using and the time they accessed the webpage. Whenever a user visits a website where the tracking code has been placed that user's data is sent to the Google Analytics system and processed. Website developers can then access this information through the Google Analytic interface, which visualizes the collected data. Since infovis tools have to map visual representations onto abstract data the visualizations created can tap into our visual abilities to find differences among visual patterns. Differences in length and two-dimensional spatial position are two properties of visual representation we as humans can immediately recognize (Few, 2009; Segel and Heer, 2010). Graphics such as bar charts and line graphs, which are used throughout the Google Analytic interface, make use of length and position to present differences among data. While there are many more examples of infovis tools that go beyond representing data using bar charts and line graphs, in general, infovis grafts a visual representation onto data which had no visual representation previously.

As I continue to describe each visualization domain many of these domains share similarities with information visualization and the distinguishing line between the domains is flexible. All visualization domains visually represent data in some way that, at the surface, can be defined as a form of information visualization. It's only when we begin to look at the other properties of infovis – providing visual representations where there are none, amplifying cognition and representing abstract data – can I begin arguing why the various visualization domains are seen as unique.

Key Attributes

- Information visualization is both interactive and computer-supported.
- Visuals are mapped onto data which is abstract and has no tangible visual representation.
- Mapped visuals are used to assist our memory and increase our comprehension of the data.
- Mapped visuals take advantage of common human perception skills such as the ability to quickly find differences between length and 2D spatial position.

Scientific Visualization

Scientific visualization is a domain associated with scientific study. It shares a number of similarities to information visualization except it relies on visualizing data with a tangible, visual form. When visualizing data that relates to abstract concepts like website traffic, as with the Google Analytics service, it is not feasible to visualize actual users visiting a website, and might look a little awkward. When visualizing data related to say the human body or the geographic locations of buildings, however, it is possible to visually render the actual, tangible properties of such data. Scientific visualization, like infovis, attempts to amplify cognition but does so through the “visual representation of scientific data that are usually [tangible] in nature, rather than abstract” (Few, 2009).

An MRI scan is an example of scientific visualization. MRI machines magnetize portions of a person’s body to align the body’s magnetic field. The magnetic field is then detectable by the MRI scanner which produces an image of the body part being scanned (similar to the type of image an X-ray machine provides). MRIs visualize the relationship between the different types of material in the body being magnetized. Bone, for example, produces a different magnetic field compared to soft tissue and is visually rendered differently in a MRI scan image. Our skeletal and tissue structures already have a visual representation in the real world, although one not easily accessed, and an MRI scan image is a visualization attempting to recreate that visual representation.

There are also cases where scientific and information visualization collide, where both tangible and abstract visual representations are used at once. fMRI scans are MRI scans of the brain which also measure blood flow through the brain. Blood flow is measured in order to determine the amount of brain activity occurring while the person being scanned is performing a task. If I am reading while my brain is being scanned areas of my brain that are associated with reading (areas that process word association for instance) are likely to have increased blood flow. fMRI scans are therefore used to associate regions of the brain to various tasks or sensory stimuli. The visualization produced from an fMRI scan is similar to a regular MRI scan. Different sections of the head are visualized according to the magnetic frequency of the tissue, which has a

tangible visual representation. Additionally, areas of the brain that experience increased or decreased blood flow are visually overlaid on top of the MRI image, often using color-coding schemes such as heat maps, representing an abstract visualization of real blood flow (I say abstract visualization because actual blood is not visualized as flowing through the brain). Increased blood flow is depicted with warmer colors like red (Figure 3.5). Scientific visualization is defined as depicting tangible data but as with the fMRI example it can be combined with abstract visual representation of data that falls within the domain of infovis.

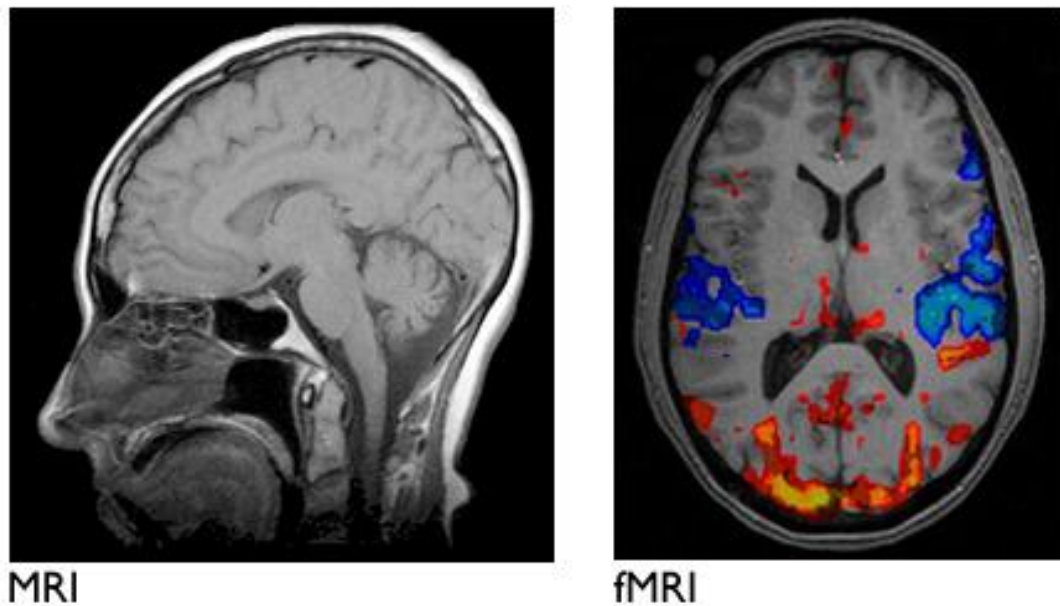


Figure 3.5: MRI (image from Hornak, 2011) and fMRI (image from Devlin, 2008) scans are visualizations of data which has a physical presence and spatial orientation, in this case bio-material in the body.

Key Attributes

- Scientific visualization attempts to accurately recreate the real world visual representations of tangible data.
- Data is often associated with scientific study but can include other areas of study such as architecture.
- Tangible and abstract data can be visualized together in order to relate one dataset to the other.

Artistic Visualization



Figure 3.6: Jason Salavon combines images of homes for sale in different residential areas across the United States in order to produce images which show common themes and patterns across the images (Salavon, 1999).

Information and scientific visualization are two popular domains and as a result visualization is often seen as being exclusively about data exploration, hypothesis forming and sensemaking (Veigas and Wattenberg, 2007). As a counter point to this way of thinking, artistic visualization is produced “by artists with the intent to make art” (Veigas and Wattenberg, 2007) and has the “explicit goal of challenging preconceptions of data and representation” (Pousman et al., 2007). Artists still take advantage of the traits that other information designers find useful in infovis, such as the ability to map data onto visual representations. However, while someone producing an infovis tool will map data onto a visual representation to promote greater data comprehension, an artistic visualization designer may use visual representations to distort data visually. Artists thus “sin” and break the rules of proper visual representation techniques (Veigas and Wattenberg, 2007). Jason Salavon “sins” when he collapses images of house reality or playboy centerfold photos into one image by averaging the pixels from the group of photos in Figure 3.6 (Salavon, 1999). The process destroys any hope of gaining information from the individual photos but in their place information detailing the

common color shades and outlines of the photos are revealed. Artistic visualization can be seen as a domain which allows visualization to “embody a forceful point of view” (Veigas and Wattenberg, 2007), giving an artist the ability to persuade the viewer intentionally. These attributes of artistic visualization are explored further when I discuss the domain of critical visualization.

Key Attributes

- Artistic visualization attempts to alter a viewer’s preconceptions of data.
- Artists routinely ‘sin’ by breaking common rules that define the “proper” methods for visually representing data.
- Visualization allows artists to present data with intent in hopes of persuading the viewer.

Minimalist Visualization

If one believes in the assumption that visualization provides analytic insight, as information and scientific visualization claim, then the main goal of visualization should be “clarity, precision and efficiency” (Tufte, 1983). Minimalist visualization is my interpretation of Edward Tufte’s work. His work promotes the idea that graphics reveal data. Tufte lays out specific edicts for designers to follow in order to “reveal data” including: avoid distortions, make large datasets coherent, encourage comparing data, reveal several levels of detail and serve a purpose. Many of these edicts agree with Few’s list of traits describing meaningful data: high volume, historical, multivariate, atomic, clean, clear, dimensionally structured, richly segmented and of known pedigree. The minimalist design philosophy seems to align itself with the ideals of information visualization while contrasting with other domains such as artistic visualization.

Tufte lays out additional rules of thumb for designing information graphics (with his early work focusing on non-computer supported, static graphics). Rules like the ‘lie factor’ of a graph – (The size of the effect shown in the graphic) divided by (the size of the effect in the data). This happens a lot when graphics attempt to create perspective, such as a bar chart growing from the ground in 3D space. The area at the top is going to

be much larger than the bottom but both sections technically cover the same distance in the data (Figure 3.7). He also warns designers against something known as chart junk (extraneous artwork, unneeded textures and overzealous grids) which draws the eye away from the data. Information graphics in newspapers often design graphics around the type of data being presented and can be considered chart junk. Housing prices are shown with houses to stand in for a bar graph, or fuel economy is shown using a road. Tufte notes, “there are better ways to portray spirits and essences [of the data] than to get them tangled up with statistical graphics” (1983) where the graphics become susceptible to the ‘lie factor’ or ‘chart junk’. Minimalist visualization is about providing as much value as possible within information graphics using clear and validated visual representations of data.

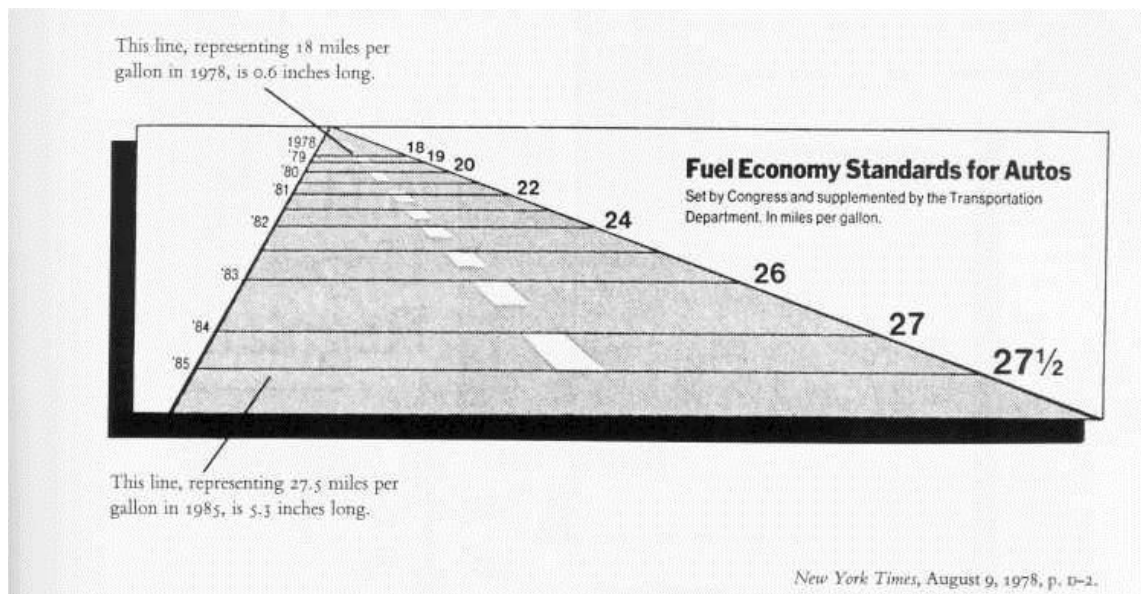


Figure 3.7: Tufte argues visualization should not distort data or add frivolous graphics. The image above does both, the 3D perspective distorts the data and the added road image is unnecessary.
(graphic from Tufte, 1983)

Key Attributes

- Visualization should never distort data for aesthetics.
- Every element of a visualization should provide value to the viewer.
- Multiple levels of detail about a dataset should be revealed in a small area.

Directed Visualization

Directed visualization concerns itself with a similar problem as minimalist visualization, the distortion of visual representation. Except distortion does not come from the addition of “chart junk” to visualization, it instead comes from the reduction of data. Directed visualization is visualization without reduction (Manovich, 2010). Manovich, the originator of directed visualization, describes information visualization as having relied on two key principles which affect directed visualization: the reduction of data and spatial variables. First, infovis systems, as Manovich argues, often remove lots of data in order to analyze a few reduced properties. Movie or game review scores are an example. Critics boil down, reduce, large media artifacts into single numbers that can be compared, leaving a large amount of data intact (such as the actors involved, the cinematograph used, the production lifespan, etc.). Second, infovis uses arbitrary spatial arrangements as opposed to scientific visualization, which uses a “fixed spatial layout of the real [tangible] object” (Manovich, 2010). Since space is arbitrary to data with no real visual layout, in a visualization should be able to manipulate other datasets we assume have a fixed spatial position.



Figure 3.8: Tag clouds provide analysis of word frequencies in a text while convey how frequently a word appears by altering the word itself (e.g.changing its size). (graphic created using Wordle.net)

Based on the two key principles of infovis, Manovich makes two proclamations in regards to directed visualization. One, directed visualization is a method of spatially organizing non-scientific data without reducing the data set. Two, media is the main source of data for directed visualization. Tag clouds are one instance of directed visualization that follows both principle rules. First, tag clouds represent the word frequency found in a text. Each word frequency does not have a set spatial location because a single word could be found throughout the entire text. A tag cloud takes advantage of this and lays out words in close proximity based on criteria such as alphabetic order (Figure 3.8). The size of each word is also altered based on the value of a word's frequency; less frequent words are smaller, more frequent words are bigger. The size differential helps users immediately see the most popular words. A tag cloud refrains from reducing the word frequencies into something like a linear table of values where frequencies could be seen but would not be in close proximity or altered in size. In this case the words represent the medium being visualized, tag clouds map properties about the words (e.g. frequencies) onto a visual representation that keeps the words intact (as opposed to something like a table which breaks up properties into rows and columns). Another example is Google Book Search (Figure 3.9) where text is the medium being visualized. When searching for a word in a book, the Google Book application provides the user with the pages where the word is found. Whole snippets of text from the book are shown with the search word highlighted. Compared to a book's index, which lists words and the page numbers where they are found, Google Book users can see the context around the searched word and make a better decision as to whether that part of the book is of interest to them. While the examples I give here are directed visualizations of text media images, videos and other forms of media can be used as well. If the principles behind Web 2.0 allowed all forms of media to become data, as I argued in chapter two, it can also be argued all media can be visualized in both a non-reduced form or a remixed form (which can remain non-reduced) using directed visualization.

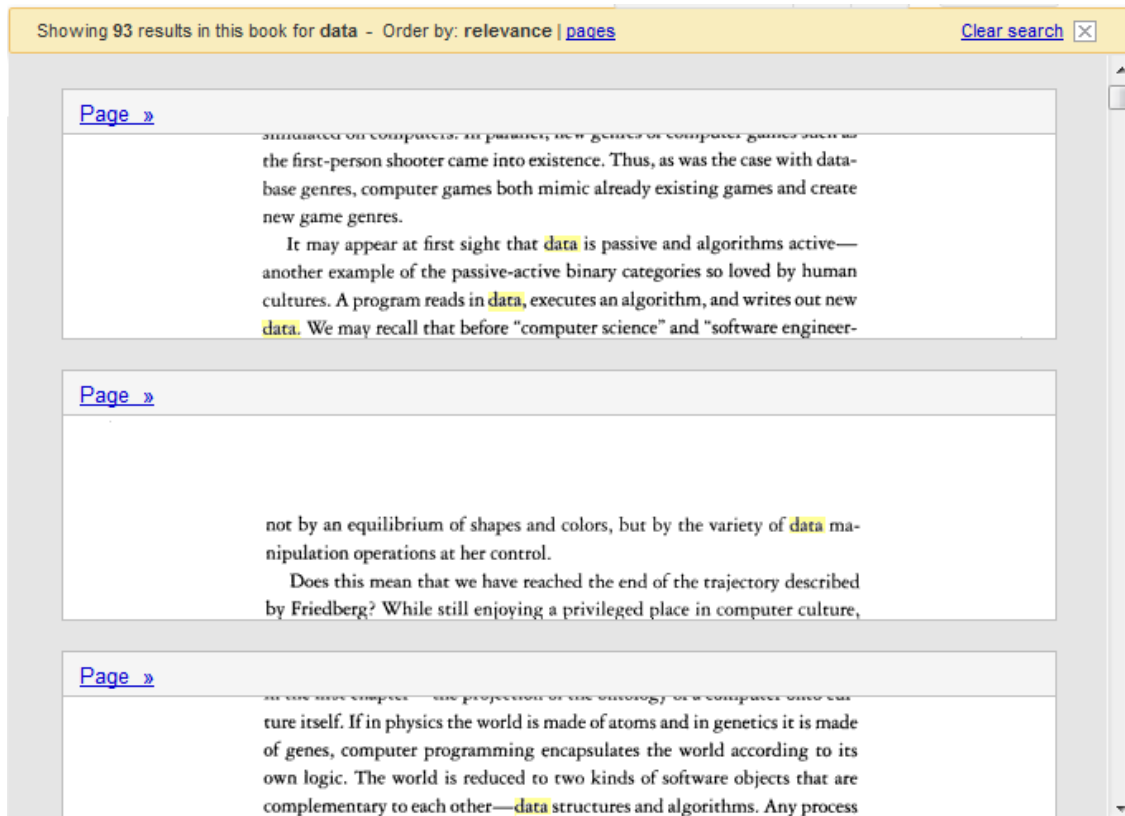


Figure 3.9: Google Book (pictured) word search provides users not with page numbers of where to find a word but actual text from the book with the search word highlighted for context.

Key Attributes

- Information visualization relies on two properties: data reduction and spatial variables.
- Data should not have to be reduced to be analyzed visually.
- Spatial variables should not always be taken as an absolute when dealing with non-tangible artifacts like digital media.
- Media (images, video, text) in non-reduced forms are the main datasets for directed visualization.

Narrative Visualization

“Data problems begin with a question” and some argues narrative is a means to answer that question (Fry, 2004; Segel and Heer, 2010). Narrative visualization does not subscribe to one definition of narrative but is a domain intended to convey stories (Segel and Heer, 2010). This domain faces a similar problem experienced by digital designers

and authors when hypertext become a possibility using digital systems (Aarseth, 1997; Murray, 1997); narrative is typically linear but interactive systems (e.g, infovis systems) hand control over to the reader and makes their experience of a narrative less predictable. With this in mind, Segel and Heer did an initial study of over fifty examples of narrative visualization in order to determine how infovis designers are using narrative as part of their visualizations (2010).

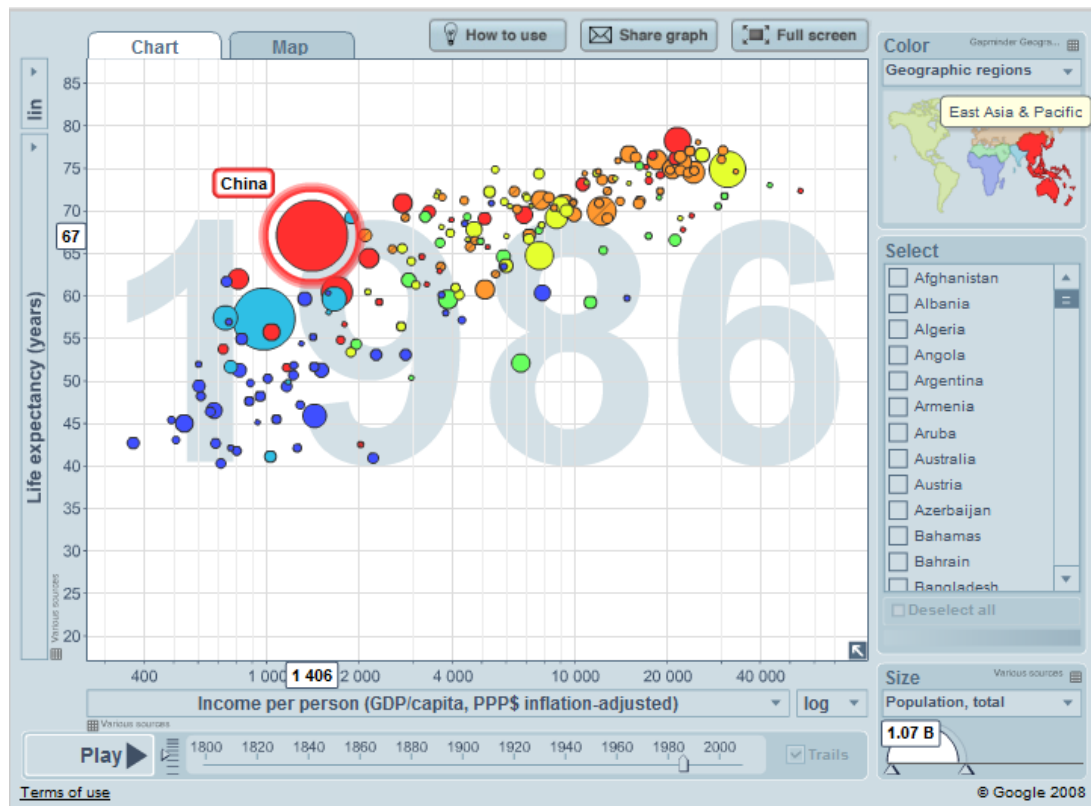


Figure 3.10: Gapminder (Rosling, 2005) is setup in a slideshow format where an interactive visualization is shown on each slide, allowing users to alter what is shown, and users can step through additional slides represented by the year the data was collected.

Segel and Heer separated the narrative structures found in the visualizations they examined into three groups: ordering, interactivity and messaging. First, order is the path a user (or viewer) takes within a visualization. Segel and Heer found visualizations can order visuals into a linear (users step through visuals in a strict order), random (users are randomly given visuals) or user-directed (users decide the order of the visuals) experience. Second, the interactivity of a visualization covers how the visuals can be manipulated by the user and how the user is taught to interact with the visuals.

Manipulation includes aspects like filtering, selecting, searching and navigation. Teaching users how to interact with the visuals include providing explicit instructions, tutorials or setting up visuals into initial configurations. Finally, messaging includes communicating observations about the visuals, adds commentary (e.g. adding labels or annotations), attaching text articles, providing introductions and adding summaries.

Following the structures found in narrative visualization, Segel and Heer concluded that visualizations can be placed between two extreme forms of narrative visualization. The first form is author-driven visualizations. These provide linear experiences with heavy-handed messages and non-interactive elements. The second form, reader-driven, is the complete opposite of the author-driven form. Reader-driven visualization have no order, no messaging but are highly interactive. Although finding a purely author-driven or reader-driven visualization is rare and the researchers give a few different examples of common patterns found between the two extreme forms. One example is an interactive slide show which pushes viewers through a linear progression of slides but within each slides viewers can interact with the visualizations being displayed. Figure 3.10 shows an example of Gapminder (Rosling, 2005), a visualization tool which makes use of interactive slides. Data visualized using Gapminder is usually historic; each year, or other set time period, marks a single slide. Pausing on each individual slide the user can mouse over data points and filter the data as they see fit. As slide progress users can watch as their filtered data changes over time. Similar to artistic visualization, narrative visualization allows an author to dictate their own views about a dataset in addition to setting up a system capable to allow viewers to create their own narratives as they explore the data, as Gapminder allows.

Key Attributes

- Narrative visualization revolves around order, interactivity and messaging.
- Ordering visuals controls how a user's experiences the visuals.
- Interactivity includes teaches users how to use features which allow them to explore a visualization.

- Messaging provides user with insightful and summary information about the data provided.
- Narrative structures and principles can be used with other visualization domains.

Critical Visualization

In his article “Critical Visualization”, Hall describes the domain of information visualization (infovis) as consisting of three parts: technology, science and art (Hall, 2008). Each part adds something different to the domain. Technologically, Hall argues, infovis is concerned with the development and selection of the best visualization solution based on the criteria of usefulness. Scientifically, infovis has a “principle” set of rules for effectively presenting data. Finally, Van Wijk argues infovis as an art form has “clear aesthetic value” and providing “intellectual and aesthetic satisfaction” (van Wijk, 2005). However, Hall does not believe the artistic aspects of infovis fall under aesthetic value alone and instead provide a valuable counter point to the technological and scientific parts of infovis as a means of reformulating data.

Hall argues that the useful and effective properties the technological and scientific aspects of infovis attempt to provide do not always work. Usefulness and effectiveness are often relative. Before the devastating hurricane Katrina that hit Louisiana in 2005 there were studies done in 2002 regarding the strength of the city’s flood-control system. As Hall describes, The New York Times created a visualization as an accompaniment to an article on the flood-control system. Hall describes the visualization as:

“A shaded relief map using a twentyfold vertical exaggeration (albeit a Tufte no-no) effectively shows the changes in terrain around New Orleans, highlighting the critical role of the levees in protecting land (shaded red) at sea level or below. A cross section of the same area reveals the water levels of the Mississippi River and ocean in relation to the land. An aerial view shows the potential path of a ‘worst-case hurricane,’ And, finally, three flooding scenarios show the city in various states of submersion. Although the case for preventive measures was

clearly and efficiently spelled out, the visualizations, like others published ahead of Hurricane Katrina, had little or no effect on policy.” (Hall, 2008, p.125)



Figure 3.11: A visualization accompanying a 2002 article detailing the problems with the New Orleans flood control system. While it is a useful visualization, the graphic had no effect on changing the policies governing flood control before Hurricane Katrina hit in 2005.

Even though the visualization effectively shows how devastating a future flood could be, a fact that was proven in 2005 after the hurricane, the visualization failed to help change governmental policy (Figure 3.11). Granted, we should not expect every graphic in a newspaper to alter governmental policy, especially if the newspaper is not even published in the state where the policy needs to change. However, that lack of expectation is the point. The 2002 graphic accurately visualized useful data (it is also arguably effective at conveying that data) but did not have the ultimate effect the creators intended. Does this make the graphic any less useful or well designed? Usefulness and effectiveness should not be the only benchmarks to evaluate visualizations, particularly when those terms cannot be seen as ubiquitously objective. As Hall argues, exploring

other, perhaps more artistic benchmarks provide other avenues for evaluating visualization beyond what they “objectively” portray.

Hall argues, the artistic side of infovis can learn from the fields of urban planning and architecture in order to become a counter measure against usefulness and effectiveness (2008). Aesthetic values are vital to urban planners and architects but the aesthetic factor is only part of their work. They “reformulate what already exists” as James Corner says (1999), including within their designs the history, local stories, economic situations, legal conditions and political interests of the surrounding area. Critical visualization should attempt to accomplish a similar feat using visual representation. It should “size up and reformulate a terrain of knowledge as well as experiment with new and alternative forms” and “array a complex combination of things that provides a framework for many different uses” (Hall, 2008). One critical visualization project is Architecture and Justice started by Eric Cadora and Laura Kurgan (2006) which looks at a correlation of crime and urban planning (Figure 3.12). A typical method of visualizing crime data is to map where certain criminal behavior takes place within urban environments. What is different about Architecture and Justice project is instead of visualizing where crimes are committed, the project visualizes data detailing where criminals live. The project found there are specific areas in a number of cities where many criminals, who are in or have been to prison, live. These areas are dubbed “million dollar blocks” because so much money is spent incarcerating the residents of those urban areas. Million dollar blocks symbolize sections of the city where the prison system is basically taking care of the residents. Cadora, Kurgan and their team use this revelation to discuss the possibility of concentrating efforts to provide these urban areas with better social services and civic infrastructure in order to curb the continued cycle of criminal behavior and prison reliance. Architecture and Justice is a critical visualization which challenges the preconception that the location of a crime is the “right” data source to explore and instead focuses our attention on where criminals live and describing how urban planning has an effect on those areas. Critical visualizations can still use the techniques forged from the technological and scientific part of infovis, creating useful

and effective visualizations but a visualization must also combine many areas of interests regarding a dataset in order to alter our perceptions.



Figure 3.12: Photos from the Architecture and Justice report which details how specific city blocks in major cities account for the highest percentages of prison inmates and discusses how these areas should be provided with better social services and community infrastructure.

Key Attributes

- Visualization has three parts: technology, science and art.
- Technology, or usefulness, and science, effectiveness, often outweigh the importance of art in discussions of visualization but these criteria can be relative.
- Visualizations should “reformulated what already exists” similar to artistic professions such as urban planning and architecture.
- Data can be visualized in relation to history, local stories, economic situations, legal conditions and political interests.

Casual Visualization

Amplifying cognition is argued as one of the main properties of information visualization (Card et al., 1999); our memory and ability to comprehend data is likely to be enhanced by infovis. This perspective portrays information visualization, specifically, as a domain helpful for a particular set of analytic tasks, useful for sensemaking, which

often revolve around work, business or other serious issues. What about using infovis in a less serious environment? If visualization is a powerful tool for visually representing data then it can likely be used in other aspects of our lives and used to gain insights beyond analytical ones. Perhaps we can use visualization more casually.



Figure 3.13: Infocanvas (Stasko et al., 2004) is an ambient visualization which maps graphics to data streams like weather, news feeds and the stock market.

Casual visualization (or casual infovis) is defined as a domain which uses “computer mediated tools to depict personally meaningful information in visual ways that support everyday users in both everyday work and non-work situations” (Pousman et al., 2007). While many visualization systems can fit this description Pousman et al. list a number of properties which solidify what it means to be casual with visualization. For one thing, ‘non-work related situations’ can reference a variety of areas. One area Pousman et al. mention is ambient visualizations that exist in the periphery of our lives, and thus these systems are built to provide information at a glance. Infocanvas (Stasko et al., 2004), for example, is a piece of software for creating abstract visualization that are digitally displayed on a monitor, which can be placed as if it is a painting on the wall or photo on a desk (Figure 3.13). The image itself contains icons or signifiers mapped to data sources such as the weather or the stock market. When the data changes, say the weather, the icon in the image changes too (e.g. the sun turns into clouds). Ambient visualizations are not intensive visualizations, requiring users to spend less time worry about gaining analytic insight and instead provides a monitoring devise users can glance at throughout the day.

Another property of casual visualization refers to the user population of a casual system as being larger and may be large but contain both experts and novices. Artistic visualization is stated as being an area related to casual visualization and can be used in situations where a diverse set of people interact with the art piece. Artifacts of the Presence Era (Veigas et al., 2004), for instance, was a museum installation that collected and manipulated sound and video recordings around the installation to create a unique amalgamation of the two (Figure 3.14). An amalgamation which the audience, i.e. anyone visiting the museum, picked up on:

“This sense of fluid, evolving time seemed to be one of the most attractive aspects of the installation to visitors. Being able to peek back at past moments in the gallery, seeing someone’s glimpse, a person’s movement, a kid’s gesture provided visitors with moments of surprise and amusement while giving them a sense of how the museum space had been inhabited in the recent past.” (Veigas et al., 2004)

The museum visitors did not have to be experts in visualization or the recording of sound/video to understand how the installation works. They went along with the visualization and found enjoyment from how it was representing the museum around them.

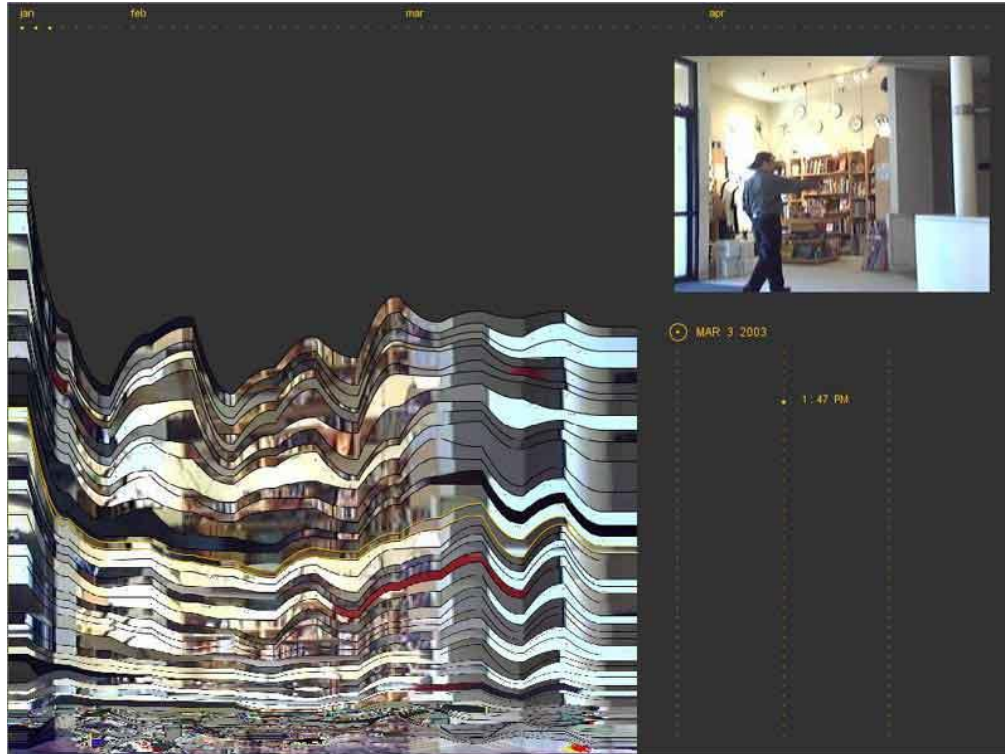


Figure 3.14: An image of the Artifacts of the Presence Era museum exhibit (Veigas et al., 2004) where audio and video recordings from the museum were combined to reflect on what was happening in the museum over time.

The third property of casual visualizations describes the data visualized in a casual system as being personally important to a user. For instance, the third type of casual visualization, social visualizations, represents data regarding friends, family and other social aspects. My personal social data is not necessarily important to other users but I, myself, may find the data personally important to casually browse. Many social network or media sharing services provide access to data relevant to single users and visualization systems are built to tap into those data feeds. Social collider (Figure 3.15) is one such system which uses twitter data from any user that is specified to the system (Schmidt and Pohflepp, 2009). The visualization produced is supposed to represent a visualization produced from a real particle accelerator/collider. Tweets act like particles that can become excited, connecting with other related tweets or spinning around, when other Twitter users responded or retweet the initial tweet. Otherwise, if the tweet goes relatively unnoticed the tweet's particle remains inert and connects to the next tweet in chronological order. Social collider does not offer advance, accurate analytic insight into

a user's tweeting behavior but instead provides an interesting visualization to browse regarding personally important data.



Figure 3.15: Social collider (Schmidt and Pohflepp, 2009) simulates a particle accelerator by using Twitter messages as particles. Messages are connected when they share similar content and messages begin to spin when other users interact with the messages.

The fourth and final property of casual visualization defines the domain as affording a number of additional insight types a person can experience when using casual visualization systems. These insights include analytic, awareness, social and reflective. Analytic insights are experience through “exploratory analysis, extrapolation, and consist of large or small eureka moments where a body of data comes into focus for a user” (Pousman et al., 2007). Awareness insights occur while consistently monitoring a data source (e.g. watching a data feed of news stories). Social insight allow users to understand their social surroundings and reflective insights provide a personal look into one's own habits or place in the world. Casual visualization therefore can provide a diverse, large population of users with the means to gain a variety of insights about personally relevant data while having less to do with work, business or more serious forms of data analysis.

Key Attributes

- Casual visualization systems can expect larger, heterogeneous user populations.
- Casual systems are built for non-work purposes often related to social, artistic or ambient visualization.
- Users often spend short bursts of time with casual visualizations but may linger for longer moments of reflection.
- Data is typically personally important to users.
- Users can gain analytic, awareness, social and reflective insights from using casual visualizations.

Personal Visualization

Personally important data is one of the key attributes of casual visualization and it is similarly a key attribute for personal visualization. The difference between the two domains is the difference between ‘personally’ important data and ‘personal’ data. Personally important data can be anything someone finds provocative, novel or otherwise useful. Personal data on the other hand is data directly related to a person, whether they find it provocative, novel or useful is another matter. Another difference between personal and casual visualization is the homogeneity of data. Social visualization is one type of casual visualization because social data is an abundant, homogeneous data source. Everyone using social networking services has friends, write public status messages and form groups. Personal data, however, is more than just social data. There are medical records, financial statements, vacation photos, education transcripts, work history, hobby interests, legal documents, etc.; each with their own data structure and each with their own level of importance to each of us. Personal data is scattered and heterogeneous; bring those data sources together presents a challenge.

Personal visualization should act as “human extensions” to our lives (Brown and Vaughan, 2009). Personal visualization systems should aid in the study of ourselves as individuals keeping a historic record of our values, activities and performance. Lifelogging, the “recording of personal life experiences by us using digital technology” (Byrne et al., 2008), is one such method where research is exploring how to record and

present data about our everyday lives using techniques such as visualization. By using simple devices such as SenseCam (a wearable sensor/digital camera) (Gemmell et al., 2006) images, temperatures, and movement within a person's environment throughout the day can be captured. Projects such as SenseCam Visual Diary (Lee et al., 2008), can visualize the images captured, allowing data to be filtered and sorted (Figure 3.16), while other projects like MyTinyBits (Gemmell et al., 2006) combine SenseCam data with other data sources including documents, presentations, and videos. Commercial systems such as Nike+ (Nike, Inc. 2010) also fall under the heading of lifelogging (Figure 3.18). While running, a Nike+ user logs data related to the run (length, position, time of day, heart rate). Users can view their run data overlaid on maps and set goals for themselves related to the data captured (e.g. I want to run 2 miles tomorrow). Using technology like SenseCam and Nike+ allows users to analyze their life in ways that normally would go unnoticed or take time to collect and review manually.

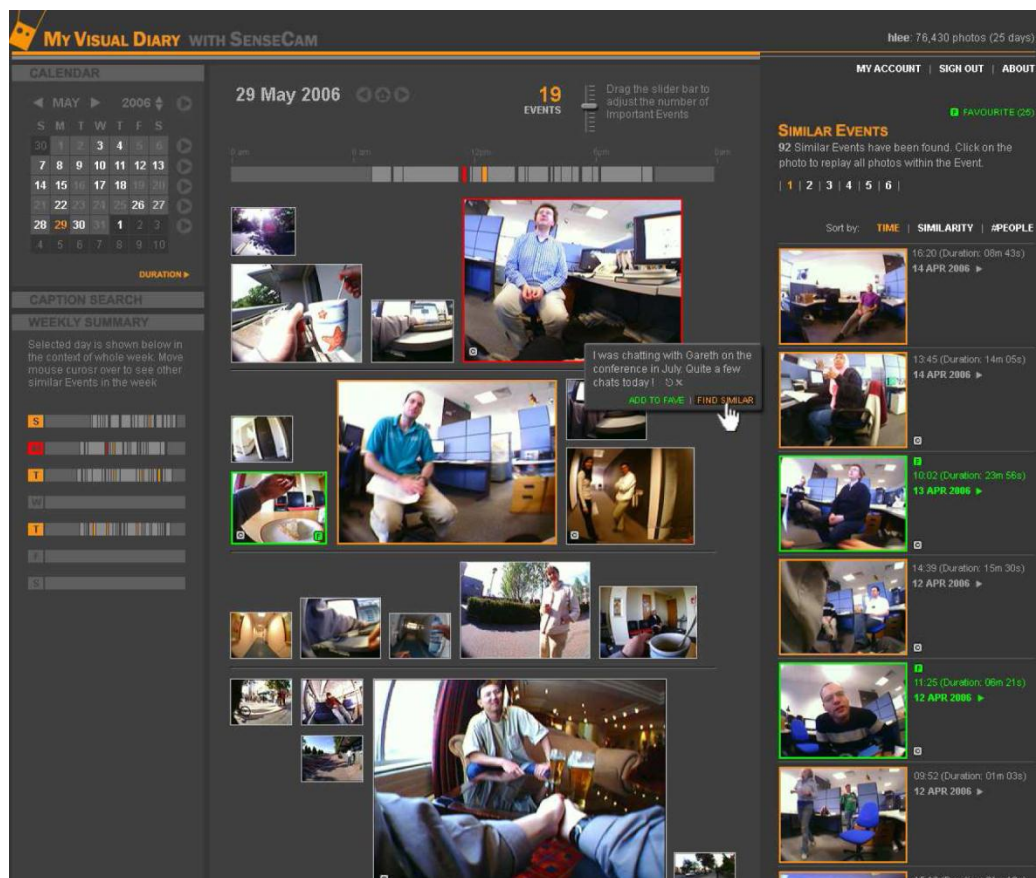


Figure 3.16: Visual Diary (Lee et al., 2008) visualizes images collected from someone's SenseCam which is wore around the neck and captures images throughout the day.

Besides the technology inspired work of lifelogging, there are examples of individuals creating personal visualizations on their own. From 2005 until 2010 Nicholas Felton, a graphic and visualization designer, created detailed self-reports about various aspects of his life (2005). He recorded the restaurants he visited, transportation he took, moods he was in and the drinks he consumed (to name a few examples). On top of this, all of the self-collected data is visualized in his personal reports (Figure 3.17). Since Felton's work has hit mainstream visualization circles there have been similar movements looking to empower individuals who have the drive to record their lives. 'Quantified Self' is one such over-arching movement which has begun holding yearly conferences and provides guides on how individuals can capture their own personal data using both hi-tech and low-fi techniques (Kelly and Wolf, 2007). Whether technology plays a major role in gathering personal data or it is simply a matter of separate individuals collecting some data on their own, the role of personal visualization is to "review, reflect and reminisce" about our personal lives while also affording the ability to analytically pick our lives apart in a functional way.

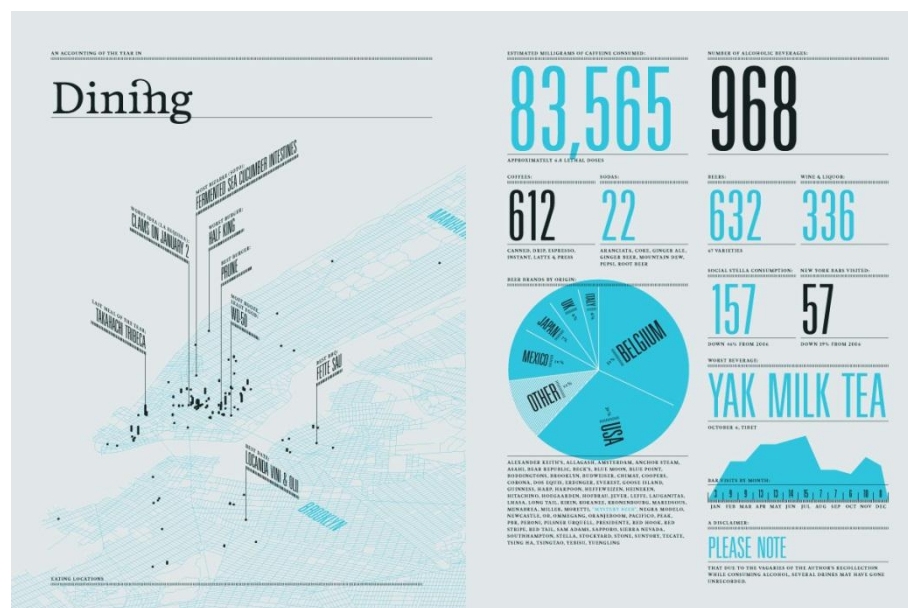


Figure 3.17: A page from one of Nicholas Felton's yearly reports (Felton, 2005) detailing events and data revolving around his life.

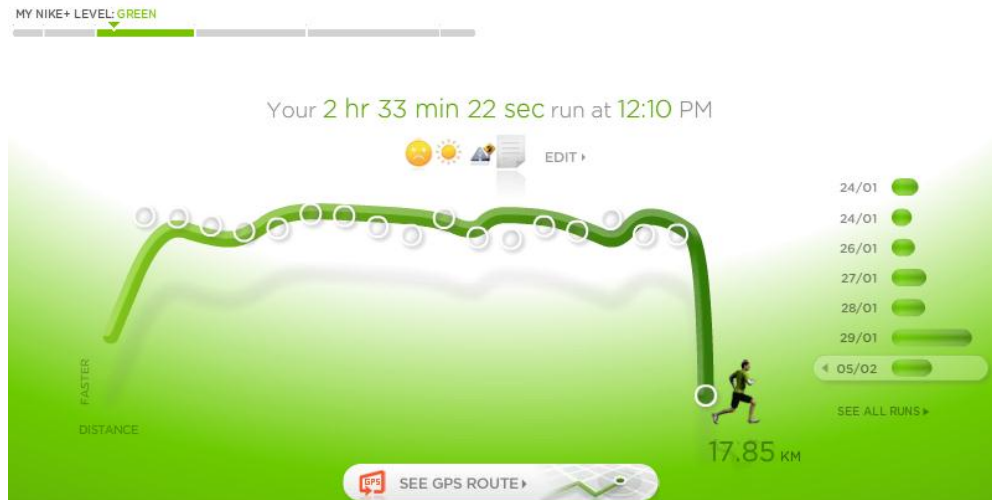


Figure 3.18: Nike+ (Nike, Inc. 2010) provides visualizations of a runner's performance over time.

Key Attributes

- Personal visualization is an extension of our lives.
- Personal data is heterogeneous and scattered.
- Data is used to review, reflect and reminisce.
- Users can use pre-built systems to capture data or collect data using low-fi methods.

Playful Visualization

In a similar way casual and personal visualization work, playful visualization is about personally important and personal data. Playing is an engaging process that requires us to invest our time and attention (Brown and Vaughan, 2009), making our play personally important to us. We create both personal experiences while playing and, typically when playing games, create our own personal play data represented as scores, avatars and replays. Playful visualization is not only about presenting data related to play so we may gain insights or reflect on our experience. Play data must fold back into our play experience, playful visualization supports but also promotes play.

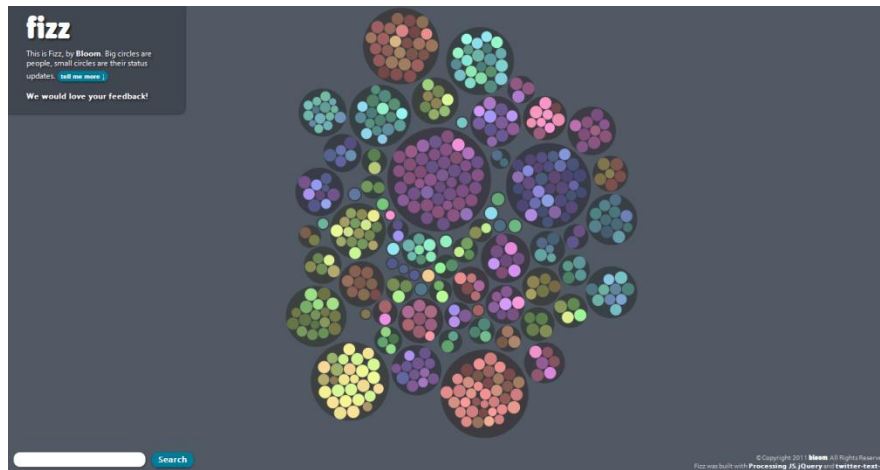


Figure 3.19: Fizz (bloom, 2011) presents Twitter messages using a physics-based visualization system where each circle represents a tweet and can be pushed around by the user.

Playful visualization can be separated into three different categories: interactive play, supporting play and creating play. Interactive play is when visualization systems are built to be playful. For example, Fizz built by the group Bloom is a tool that visualizes a user's twitter feed (bloom, 2011). Tweets from users being followed pop on the screen as little circles and clump together if they are from the same person (Figure 3.19). Users can mouse over each circle to see the tweet and can grab the circles to move them around, which collide with other circles in the process. Fizz is certainly not the most streamlined way to view a twitter feed but promotes exploration while providing useful information such as which follower writes the most tweets. Another interactive play project is called Kimono Fishing, built by Colleen Macklin and her students. A user is given a “lure” that represents a traditional ingredient used in kimono production for making color dyes. The “fish” are the historic colors produced from the possible ingredient lures. When a user has an ingredient lure selected the different colors that are produced from that specific ingredient start moving towards the user's mouse cursor, eventually attracting the related data to the user. Interacting play utilizes concepts like collision, movement and attraction to visualize data.



Figure 3.20: Halo waypoint provides Halo players with support information such as how they are progressing in the game.

Second, visualizations that support play provide additional means for approaching playful activities. One supportive play system for playing a game is Microsoft's Halo statistics website called Waypoint (the successor to Bungie.net). Every online match a Halo player has played in the past is tracked and displayed on the website (Figure 3.20). Overall achievements, ranks, game history are all recorded. Every match is provided in great detail describing how each player, or their team, progressed through the match and the final overall results. Players use these systems not only as trophy rooms, where their achievements are housed but can also gain insights into their play behavior, and potentially get better at playing the game.

Another supportive play system is Cyclopath (Figure 3.21), a visual tool for mapping bike routes around the city of Minneapolis-Saint Paul. Cyclopath's user population is crowd-sourced in order to mark valid bike routes in the city and to provide changes to the roadways on the overall map. Other mapping tools like Google Maps do not have a robust method of altering a map's structure, nor has the ability to mark routes that do not conform to the roadways. When attempting to find a bike route in Cyclopath multiple different routes appear. Some take cyclists down major streets, some through

parks and other show shortcuts that may have gone unnoticed if not for another user adding the route to the system. Cyclopath supports cycling, which can be a leisure or playful activity, by allow a community of cyclists to share their experiences within the city of Minneapolis-Saint Paul.

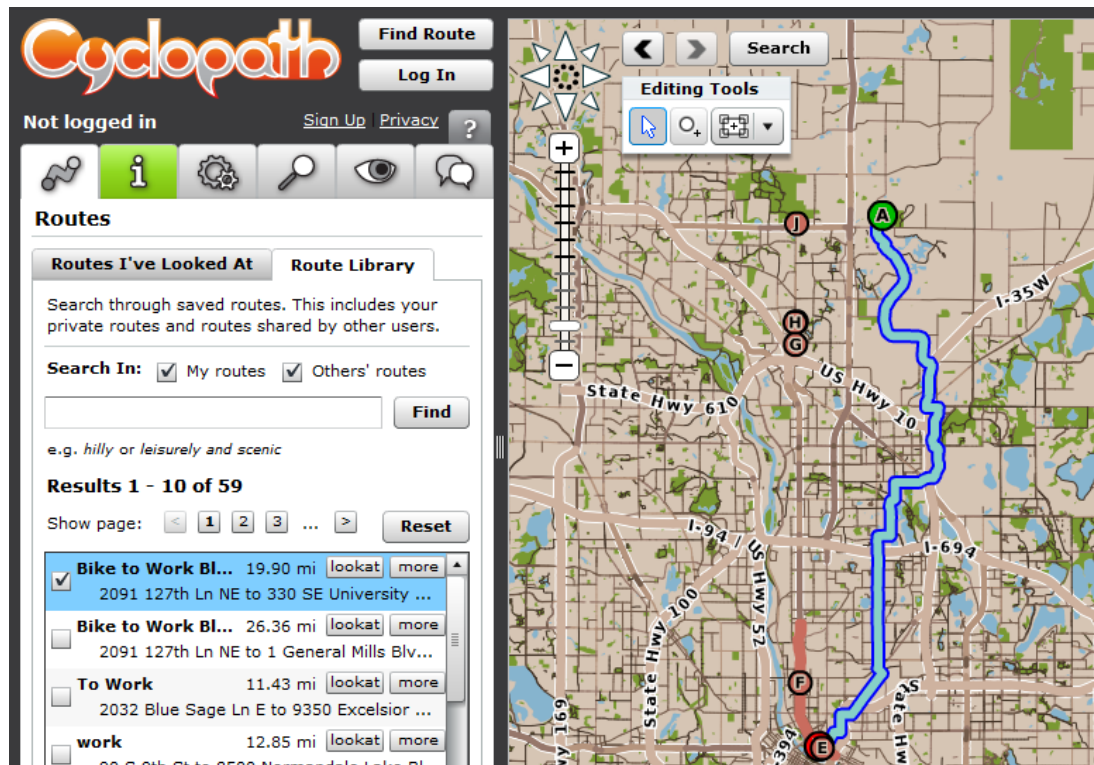


Figure 3.21: Cyclopath is a mapping tool for planning and viewing bike routes in Minneapolis-Saint Paul, MN.

Finally, creating play refers to users who find creating visualizations to be a playful activity. This means players and spectators make use of data feeds related to playful activities to build their own visualization systems. The game Spore, for example, provides data feeds for players to use which relate to the game's creator creation tools (which is a 3D model editing program built into the game and has a major role as part of the game's mechanics). One part of the game allows players to create both their own alien creatures and other inanimate objects (space ships and buildings) using the in-game creation tools. As a result players have produced millions of creator and object assets using the creation tools. Players can access all of these assets using the Spore API (application protocol language), which contains data regarding how to construct the player created models. This spawned a number of Spore creator visualization programs.

One of which was Spore Skeleton (Meyers, 2009), a program built to display the underlying skeleton of the Spore creator models (Figure 3.22). Normally this data is hidden from the players and is only useful for the game's engine to use. Having access to the Spore API data, allows players to create systems which the game developers may never build but are playful for players to build and use.

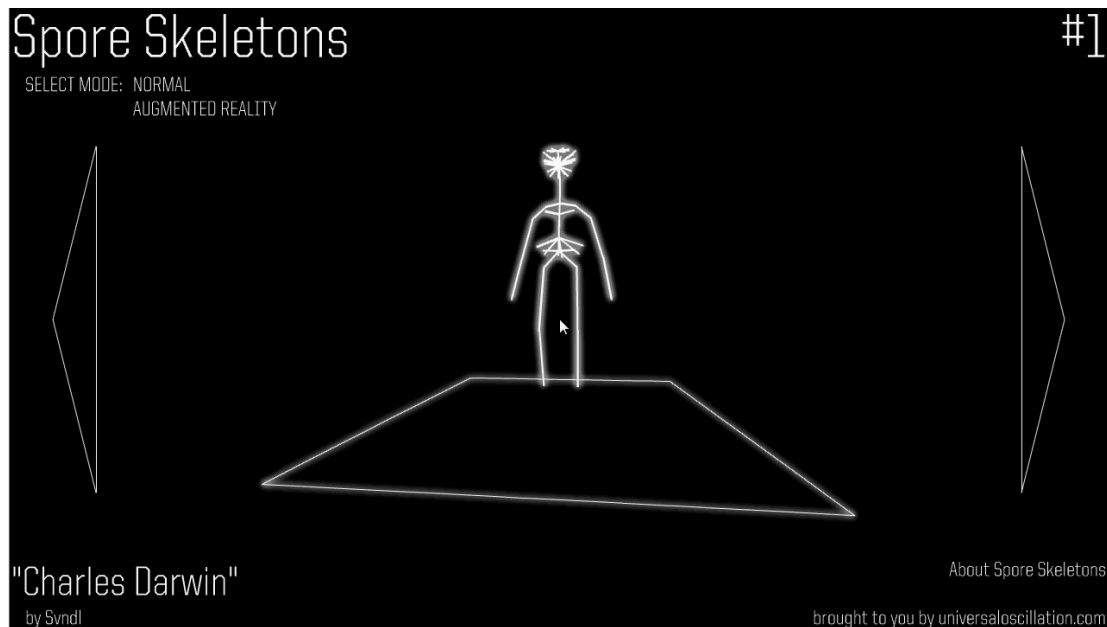


Figure 3.22: Spore Skeleton (Meyers, 2009) was built using the Spore API which allows players to build their own visualization using data available from the game Spore.

Key Attributes

- Playful visualization should support and promote playful activities.
- Visualization can provide playful interactive visualizations which draw in users to play with the visualized data.
- Supporting play requires visualizations to provide value to users in a way that can be taken back into their play experience,
- Creating visualizations should be seen as a form of play which can be promoted by data providers by making data easier to access and interesting to visualize.

Summarizing the Visualization Domains

After reviewing the visualization domains presented we can begin to see common threads running through the domains. Certainly visual representation is a big issue for

some of the domains. Minimalist and directed visualization are almost entirely devoted to arguing for specific procedures to visually representing data (which is why I designate them visual representation and design philosophies in Figure 3.3). The arguments made by both minimalist and directed visualization can be applied to the other domains I presented, including both the major visualization domains and the sub-domains. Some domains are concerned with the message visualization designers attempt to convey. Narrative, critical and artistic visualization all speak about creating messages through stories, putting data into context and challenging preconceptions. Then there are the domains that focus on the personal importance of visualization for the user. Casual visualization speaks about using visualization for everyday activities, personal visualization attempts to enhance our understanding of ourselves and playful visualization promotes using data as part of playful activities. What these threads point to is the duality that exists amongst the domains of visualization, one that addresses a functional aspect of visualization (e.g. proper visual representation methods and the curation of messages) and one that address the experience of visualization (e.g. a contextual and personalized endeavor). I explore this duality next as I apply the functional and phenomenological framework to the visualization domains I have presented thus far.

The Functional and Phenomenological Aspects of Visualization

A visualization system can be interpreted as a type of categorization system and as a system experienced by users. Each of the example visualization systems presented make data visible, compare data through relations, control how data is compared and control how data is visually represented. From the user experience perspective, users experience a sense of exposure as certain datasets are visually presented. Sometimes the visualized data is personal making the exposure of the data different compared to exposing data related to impersonal areas like census data or the weather. Visualizing relations between data create interconnections we may have missed or otherwise are made more aware of through the system. Increasing “sensemaking” through connecting data is a common goal for domains like information visualization but user can also become interconnected with other people (through casual visualization for example) or

their personal environment (as critical visualization tries to accomplish). Finally, temporal data is routinely used by visualization domains and create a history for users to explore. Sometimes this history can be abstract or artistic, as we saw in project Artifacts of the Presence Era (Veigas et al. 2004), other times empirical, similar to how Gapminder presents demographic data. Exploring each set of functional and phenomenological properties in succession below, I present how the visualization domains, and their systems, are relate within each property and where they differ.

Visible

The term visualization implies data is made visible. As to the question of what type of data is made visible, most of the visualization domains are used to visualize non-specific datasets. There is not an arbitrary limit on what information visualization or critical visualization, or many of the other domains, can visualize. Only personal visualization is used to represent a specific dataset (personal data) and there are other examples of visualization domains focusing on a certain type of data, for example geographic visualization visualizes geographic data (MacEachren et al. 1998).

The difference between the visualization domains in regards to making data visible is how each domain represents data. Scientific visualization is used to represent data with a tangible and spatial visual representation. Information visualization must map a visual representation onto abstract data. Artistic and critical visualization attempts to represent data in a provocative or antagonizing manner in order to engage a visualization's audience. Narrative visualization represents data by adding a storytelling component for a user to traverse. Finally, minimalist and directed visualization are the two domains that discuss data representation directly. The representation principles minimalist visualization proposes argue that visuals need to provide value to users and should not be cluttered with extraneous graphics. Directed visualization argues that data does not need to be reduced before it is visually represented. Data such as images and video can be visually represented with minimal reduction according to directed visualization. Even though the visualization domains can, with some limitations, be used

to visualize any dataset, how they set about representing the dataset differentiates the domains.

Last, visible data within a visualization routinely has a relationship with the visualization's audience. This is not a relation or comparison between data, but how the data itself relates to a user's life. Casual visualization, for example, represents data that is personally important to a user; data a user wants to explore such as data related to their social circles. At the same time casual visualization represent datasets that do not require consistent attention or require the user to perform intense analytical data analysis. Casual visualization are structured to be casual, hence the visualization provide users with something other than analytic insights alone. Personal visualization is presented in a similar way. Personal data is personal to a specific user, making the data highly relevant (personally important) to a single user but often meaningless to other users. Personal data can be represented as a means to analytically investigate one's life but, like casual visualization, can be used to reflect of reminisce as well. Data used in a playful visualization, while not limited to personal data or any other specific type of data, can present data casually and often provides users with personal data related to their play. Playful interactive visualizations are meant to be played with and do not require users to approach analyzing data in a normative fashion. In contrast, domains like information and scientific visualization are argued as approaching analytics as primarily for intensive, serious data analysis (Thomas and Cook, 2005). A visualization's relation to the user in these domains is a business relationship, built for pattern recognition and sensemaking in order to influence future decisions. Users need to gain analytic insight from information visualization rather than have a non-normative experience as I argue playful visualization can provide.

Visualization domains decide what data to make visible based on the data's subject matter and the relation it has with the user or audience. These are factors that decide what a dataset provides to an audience. A dataset provides worth to a user if the data is related to a personally important area of interest (whether for work or other interest) or is personal (making the dataset specific to a single user). The visual

representation of the data is used to define how data is visibly presented. This means visualization developers have to decide if their datasets will be presented using a narrative, abstract shapes, a tangible spatial representation, a non-reduced format, or any other visual representation discussed as part of a visualization domain. Deciding what data to visualize and how the data is made visible are the first steps towards enriching data using visualization. The next step is relating or compare data together so information can form.

Comparability

After data is made visible for a visualization to be complete the data must be compared. This dissertation does not go into detail about the numerous types of analysis or formulas used to relate and compare data. A dissertation on data mining or data analysis techniques would be better suited for that task. However, given the descriptions of the visualization domains, there are five different analysis categories used to compare and relate data I wish to highlight.

First, there is comparing data analytically. Pousman et al. argue that information and scientific visualization are often used to compare data analytically and to gain analytic insights. These insights refer to observations (patterns, trends) which are valuable to the field or source where the data is collected. Using a tool like Google Analytics, a website developer can correlate the location of a website's users with how much time they spend on the website. This can lead to an analytic observation that describes which parts of the world the website has penetrated and where advertising may have great impact. Observations often lead to hypothesis, need to be verified as correct and explain the patterns found in a dataset (Saraiya et al., 2005). One of the definitions of information visualization is to help amplify the cognition of a user (Card et al. 1999), which essentially means to increase someone's comprehension of a dataset and lead to valuable observations. Many of the visualization domains stated in this chapter can be approached with an analytical mindset, attempting to gain comprehension and insights from a system, even if those same visualization domains also afford the other data analysis methods I state next (for example, personal visualization can be analytical and

promote reflection and reminiscing). Comparing data analytically means systematically approaching data to gain comprehension and observe patterns in the dataset to develop further hypothesis, which can lead to further analytical data analysis.

Second, data can be monitored. Unlike comparing data analytically, monitoring data means watching a constant stream of data for a particular event to occur or reviewing new data. One example of a visualization that uses monitoring as the main source of comparing data is ambient visualizations, one type of casual visualization. The ambient visualization project Infocanvas, visually represents data for users to monitor throughout the day but does not require a user to keep a consistent eye on the data over time, nor does it make a user intensely study the data. Sources of data like the stock market, weather, email, news or scheduled events are visualized using Infocanvas. Monitoring can also be used for artistic or playful visualizations too. The ‘Artifacts of the Presence Era’ project allows museum goers to monitor the images and sounds taken from the museum over time. Systems like Nike+ or Waypoint allow runners and players to monitor their performance over time, either as a means to monitor their health or monitoring their play, respectively. Monitoring, therefore, does not compare data for users to gain observations or analytic insight. It compares data based on a specific type of data feed and certain criteria a user is interested in, notifying users when specific situations occurs.

The third way data can be compared is through reflection. Reflecting on data questions the visualization process. This means a user reflecting on data questions what the data means, how the data is represented and how data analysis is used. Given the attributes of critical visualization, it is one of the visualization domains that centers around reflecting on data. Critical visualization seeks to “reformulate what already exists” to rethink and question how data is visualized. Data is not analyzed as an objective representation of our environment, something that analytical data analysis or monitoring data analysis methods routinely do, but instead see data within a larger body of history, local stories, economic situations, legal conditions and political interests. Similarly, artistic visualization seeks to “alter a viewer’s preconceptions” to force an

audience to look at data from a different perspective. Spore Skeletons, although referenced under the playful visualization domain, is an artistic visualization that forces its audience to look at the underlying skeletons from Spore models instead of presenting the normal models which include textures and color. The project alters a player's perspective of a game model away from the outward appearance and towards an appearance the game's engine sees (a model's skeleton determines how a model animates within a game engine). Reflecting on data means data is analyzed through a holistic process attempting to approach a dataset from multiple perspectives and within multiple contexts.

Reminiscing is the fourth way to compare data. Reminiscing involves provoking past memories and experiences through data analysis. Lifelogging, a practice often associated with personal visualization, is an example of how data analysis is used to reminisce about personal life experiences. On a daily basis, image, sound, video and biometric data can be captured and presented to users for reminiscing about that day's events. User can then return to the data over time, using lifelogging as a method of reliving their past days. Lifelogging can be seen as combining the principles of directed visualization with personal visualization. Personal data is used to represent a user's daily experience and that data is presented in a non-reduced format. Non-reduced data retains a much of the original sensory data collected to provide the best chance for users to remember the experiences represented in the data. Reminiscing is therefore about comparing data in order to remember past experiences, not necessarily to question the data or gain additional insights.

The fifth and final way data is compared is through creation. While comparing data can be a static process, such as viewing data on a static Cartesian graph looking for a pattern, there are also many examples of interactive, modifiable systems which allow users to creating their own data relations. For instance, one type of playful visualization is creation play where a user will take data and create a visualization system as a hobby or personal project. The Spore Skeletons project was created by a user as a personal project without any major intention behind presenting the data from the Spore game. Personal

visualization projects like Nicholas Felton's work is an example of how people can find it enjoyable to produce their own personal visualization using data they capture themselves. Finally, even if a system is created to compare data analytically, many general visualization tools like Tableau and Microstrategy allow users to create their own visualization and reports. While the purpose may be to analyze data, the process through which users and developers create the system to compare data can have an impact on how data is compared. Some users may find the creation process the most informative.

While there are many methods for comparing data, the five types I stated above relate to the reasons why data is compared. Example of how data is compared can fit within the five type. Using statistical methods of analysis can fit under analytical data analysis. Comparing photo based on meta-data could be considered a type of reminiscing analysis. Aggregating news stories based on topics can be seen as a method of monitoring data. Organizing methods of comparison based on how the user can approach the data provides a way to discuss why one form of analysis may create a different experience when compared to another.

Control

Having control over the development of a visualization system (or categorization system) means determining what data is made visible and how that data is compared. Given the ten visualization domains covered, having control over how data is visible and compared refer to three different factors. First, the visual representation of data is controlled by a visualization creator. Second, the message, or messages, of a visualization are controlled and presented to an audience. Third, control of a dataset, or a visualization, may be handed over to an audience allowing them to manage the dataset on their own.

Controlling visual representation means determine how to visually present data made visible by a visualization system. The minimalist and directed visualization domains, for example, are specifically about controlling how data should be visualized. Minimalist visualization practices say never to distort data, create visuals with minimum amount of "ink" or chart junk, and provide multiple levels of detail. Directed

visualization suggests data, particularly media artifacts, can remain unreduced giving an audience the chance to analyze data as a holistic artifact rather than a set of attributes. In contrast, other domains like artistic visualization may find distorting or reducing data are valid methods for altering an audience's perspective, presenting data visually from a different angle. While there techniques differ, each of these domains detail methods visualization designers can use for controlling how data is visually represented.

The choice a designer makes to visually represent data ultimately has an effect on how an audience experiences the data. If the Artifacts of the Presence Era project had followed the methods laid out in minimalist visualization the datasets containing the images and sounds from the museum would have never been combined with each other, forming the layers of distorted images. However, the question whether a non-distorted version of Artifacts of the Presence Era would be novel enough for museum goers to find the project interesting still remains. If a tag cloud, which is a type of visualization that uses directed visualization principles, followed the principles of minimalist visualization it would be a list of words from a piece of text next to their frequency within the text. Having a list of frequencies would benefit someone studying a piece of text analytically but a tag cloud, instead, can benefit someone reflecting on or reminiscing about the text. Perhaps by distorting data in the Artifacts of the Presence Era project also offers an audience the ability to reflect on data rather than forcing them to analytically analyzing the data, as the principles of minimalist visualization often accommodate. Altering visual representation, therefore, means controlling how an audience comes into contact with a dataset and gives them clues regarding how they should approach analyzing the data.

While visualization systems are often said to be “interactive” and provide valuable data to an audience (Card et al., 1999), they also provide the means to control the dissemination of targeted messages, similar to other mediums like film or print. Architecture and Justice (a critical visualization) and the Felton reports (personal visualizations) are examples of visualization projects with a large amount of control over what data is collected, how it is visualized and how it is presented to an audience. These are author-driven visualizations (using Segel and Heer's definition from narrative

visualization (2010)) where the author holds most of the control over how data is disseminated to an audience. Similarly, other systems like Google Analytics (an information visualization) or Gapminder (a narrative visualization) allow an audience to explore a dataset but generally control the visual layouts, deciding what data is important to deliver to an audience upfront even if the audience can dive deeper into a dataset. Controlling the message of a visualization takes control away from an audience in an attempt to convince the audience that the message, or visual display of the data, is correct.

Finally, visualization designers do not need to control every aspect of the visual data analysis process and can give control to the audience themselves. Designer may wish to do this when the audience finds a particular dataset personally important or if the data is personal to a specific user. Bungie.net has in the past released APIs that allow Halo players to access their data from the website. Their data is visualized on the Bungie.net website but with an API players can create their own visualizations. Giving an audience more control does mean taking away a curated, controlled visual presentation of the data but it gives an audience the ability to create their own visualization systems. Spore Skeleton (a playful visualization) was built using the Spore API and most likely would have never been built by the creators of Spore. Likewise, Social Collider (a casual visualization) would not be possible without Twitter providing an API to access twitter messages. APIs are one example of aiding an audience to create their own visualization and analyze their own personal data. The API creators therefore still control some aspects of the visualization process (what data is collected, who can access the data) but they need not control every aspect of the visualization process.

Exposure

As a visualization audience we experience data as an exposure. A visualization creator decides what data to collect, how to store it, how to transform it and how to visually represent it. Eventually, the data must be revealed to the outside, to be experienced by those seeking the relations within the visualization. A visualization domain like information visualization seeks to expose data in order to assist our memory

and increase our comprehension of the data. Scientific and minimalist visualization has a similar goal. Designers within these domains attempt to create an experience where an audience finds the data visualized useful. An audience is supposed to use an information or scientific visualization to uncover insights that will affect their some important aspect of their work or life. Whether or not these tools actually achieve creating this experience is another matter.

Other visualization domains question how data should be exposed and if the right data is being exposed. Directed visualization, for instance, believes information visualization often reduces data to an extreme degree, cutting off the ability for an audience to gain useful information from the data. By showing data, particularly media data (films, artwork, images), in a non-reduced form, directed visualization is argued as creating a more fulfilling experience by exposing more data to an audience. Those who argue for critical visualization, similarly, believe visualization should create an experience to question and reformulate the data currently available. Instead of arguing data is often reduced, critical visualization creates an experience to question why certain data is used over other data. Critical visualization questions why data is revealed, but not necessarily to reveal more data (as directed visualization often does). The Architecture and Justice project visualizes data in a way that could work for visualizing crime data based on where the crime took place. Instead, it visualizes where the criminals life, taking away the location of the crime and replacing it with the criminal's resident. The project alters the data expose to give an audience an experience of questioning how we expose data related to crime. Whether it is information visualization revealing data to be useful, directed visualization revealing non-reduced data to provide a holistic analysis or critical visualization revealing data to question what data is exposed, how a visualization domain seeks to reveal data affects how an audience experiences the exposed data.

Additionally, privacy is one concept associated with interpreting the exposure of data because exposed data often has some relation to a person or group of people. When data about a person or group is exposed it reveals something and that revelation has the possibility of turning data into publically available information. The Artifacts of the

Presence Era project, for example, reveals the sights and sounds of a museum over long periods of time. Museum goers interacting with the project become a part of the data being exposed. Any museum goer in the future can then reference the past data taken from the museum. If visualization has an effect on how an audience experiences data, then it has the ability to take data related to a person and give a context that may not reflect the person completely. In essence, visualization can create an identity for a person, by exposing select pieces of data, which does not reflect the true, holistic, identity of the person. The issue of privacy arises when data that seemingly represents the identity of a person misrepresents the person or others misinterpret the exposed data from that person. As Boyd describes, “people must trust their interpretation of the context, including the people in the room and the architecture that defines the setting” (Boyd, 2010). If someone cannot trust how an architecture, the visualization systems, exposes data and the people using the architecture to interpret the data within the context it was gathered, then an issue of data privacy arises. Experiencing data through a visualization, therefore, is not only about experiencing what data is exposed, or why it is exposed, but includes an experience of how a visualization represents us when exposing data.

Interconnection

The interconnections experienced within a visualization system are based on which data analysis methods the system affords. For instance, a system like Google Analytics affords the ability to monitor traffic coming into a website. A developer can begin to see the interconnections between the new content on the website and increased traffic. The developer may then decide to create similar content in order to increase traffic further. Using Google Analytics to monitor traffic is a type of ‘ready-at-hand’ experience, the average experience that Google Analytics affords. If traffic is going up or stable, a web developer can assume they are making good decisions. If a website’s traffic starts to decrease, however, a developer has to switch from monitoring their traffic to analyzing, and perhaps reflecting, on their website traffic data. The developer enters into a ‘present-at-hand’ experience, because the normal method of monitoring traffic has failed. They need to determine why they are losing traffic and what the cause may be. Fortunately for a developer, Google Analytics affords the ability to analytically analyze,

or reflect on, captured tracking data from a website, just as it affords the ability to monitor the data. Although, it is the ability for a user to switch between these affordances that defines the experience a developer has with a visualization system like Google Analytics.

Even though Google Analytics affords the monitoring, analytic and reflective data analysis methods a user may use one or all three at any time. I argue that each of the five data analysis method describe in the comparability section above, including the three Google Analytics affords, relate to Heidegger's concept of moods. A visualization system, therefore, affords different analysis moods, which users can fade in and out from depending on what they are trying to accomplish or how they are feeling. This allows users to form different interconnections while using a visualization system. Someone using Social Collider to view their twitter messages may approach the system in a reminiscent mood (a mood that Social Collider can be argued as affording), where they seek to remember how their old tweets related to their followers or particular friends. But some users could approach Social Collider in an analytic mood (a mood Social Collider is not necessarily built for), wanting to gain insight regarding which messages they sent had the most impact amongst their followers. Even though a system may not be built to handle certain data analysis moods, a user can still alter what the system affords and change how data interconnects within the system.

Finally, experiencing interconnections within a visualization is also a process of connecting to other users and to the data itself. Users form communities around visualization system or the visualization aid communities that already existed. Cyclopath has a community of cyclists in Minneapolis-Saint Paul that use the system to share their common routes within the city. Of course the system makes these routes 'visible' and 'compares' routes based on a user's search criteria but users can create discussions about routes as well. Users connect with each other and analyze the routes visualized together. Other systems like Nike+ allows similar community features where users can share their data with friends. However, interconnection doesn't just mean using data as a means of connection users it can be connecting users to data as well. Bungie.net for example

connects users to their medals they have been awarded in various Halo games when they complete certain tasks or criteria. Certain metals carry with them more weight (such as the ‘kill 1000 opponents in a row without dying’ metal) and players can make assumptions about a player’s ability based on those awards. Even though metals are just data made ‘visible’ by the system, the value the Halo community places on the metals defines how the metals connect to the players. Thus, interconnection can be used to describe how users experience each other within a system and how visualized data connects directly to them.

Historic

The historic interpretation of data and visualization can be approached in two ways. First, a visualization is historic when it has a temporal dimension. Having a temporal dimension means data can be assessed over time and anomalies within a time frame can be explored. Many visualization systems previously discussed are built on the fact that data changes over time and those changes are important to acknowledge. Gapminder, in particular, reveals country-centric data over many decades worth of time and helps present changes related to the health and wealth of each country; changes that can be correlated to events within a country’s history that may explain why the changes have occurred (e.g. new medicines increase the life span for a country’s population). Artifacts of the Presence Era also presents data across time, visualizing the sights and sounds of a museum throughout the day. As data is collected over time, users can examine older records from the previous day, witness the museum at night or see how active the museum was yesterday. For example, the system produces dark, quiet bands of time when the museum is closed. In other words, using time as a dimension allows a system to highlight prejudices affecting the dataset over time. The users examining the interface for Artifacts of the Presence Era can see at night the museum is dark and quite, or Gapminder users can see the eradication of a disease over time after the invention of the proper vaccine. Visualizations can help find interesting temporal events that altered a dataset, and those events represent prejudices an analyst needs to know.

Second, the historic property can be used to investigate the historic prejudices of a visualization system itself. Relying on a visualization to provide an objective approach to analyzing historic data should be highly suspect, as much as the concept of historical objectivism, which Gadamer argues against in his work on historical interpretation. Having data with a temporal dimension does not necessarily clear a visualization system of the prejudices by which the system is built. For example, as part of Few's list of traits describing meaningful data the trait 'historic' is one such trait. He mentions that "historical data should be consistent or adjusted so that it is comparable over time if record-keeping conventions have changed" (Few, 2009). Data, as Few argues, should therefore be altered to meet the current standards required to study the data, throwing out any in-consistencies that may appear. Perhaps some data was lost one year and can be swept away by averaging the two points of data on either side of the missing data. "Cleaning" data (i.e. altering or reformulating data) for a new visualization system means imparting a set of prejudices on the data to make the data work within the new system. Prejudices ingrained within a system are harder to find and assess in regards to the type of prejudices found while analyzing historic data (i.e. the first approach to interpreting visualizations historically I discussed). Visualization developers do not always release the methods they use to produce a visualization and therefore ingrained prejudices can stay hidden. The task for finding and revealing a system's prejudices may then fall to outside parties, the public or those studying visualization systems. However, finding ingrained prejudices may become even harder as a system ages. Data being constantly manipulated and new features being added over time can shroud the underlying system behind a visualization has worked over many years.

Summary

The purpose of this chapter is to show how diverse the topic of visualization has become. There are a few major domains that visualization research and projects typically fall under (information, scientific and artistic visualization) but there are also a number of sub-domains that expand upon these major domains (I cover casual, critical, narrative, personal and playful visualization in this chapter but I could have covered more). There are also visual representation and design philosophies like minimalist and directed

visualization that argue for certain aesthetics and formats for visualizing data. If all of these visualization domains exist and are used to visualize data to an audience we can assume that if play analytics may take advantage of any of these domains, given that play analytics systems often visualize data as a means of disseminated game-related data to players.

I also showed that my functional and phenomenological framework could be used to analyze the presented visualization domains too. It is my hope that the analysis of the visualization domains using the framework in this chapter will set up my larger analysis I perform in chapter five, when I cover the content analysis I performed as part of this dissertation. Additionally, I listed a number of ‘types of data analysis methods’ in the compare section of my functional/phenomenological analysis above. The five types I present – analytical, reflective, monitoring, reminiscing and creation – are used in chapter five when I analyze the play analytic systems from my content analysis. These five data analysis methods play a fairly significant role in later chapters and I use them to describe what analysis methods play analytic systems afford and what methods they lack overall.

CHAPTER 4

ANALYZING GAME ANALYTICS

With data and visualization covered in the last two chapters, I now turn my attention to combining these concepts with games. As mentioned in chapter two there are many examples of game-related data. Everything from a game's content to a game's sales figures can be valuable data to different audiences (whether that audience is a game developer or an outside party like players). Visualization is also a practice used often to represent game-related data. Visual representation is both used inside of games for representing data related to gameplay (for example, features like mini-maps found in real-time strategy games) and used outside of games for representing data such as player variables or game events (as the Bungie.net system in chapter three described). This chapter combines the three terms – data, visualization and games – in order to define the concept of analytics and how analytics, which is often described as the study of analysis, is used in relation to games.

With so many different types of game data and visualization domains to discuss, I have yet to present one concept that approaches data and visualization from a larger scope. This concept known as visual analytics is an attempt at a holistic approach to data analysis. Thomas and Cook define visual analytics as the “science of analytical reasoning facilitated by interactive visual interfaces” (2005) and seek to incorporate and extend information visualization practices (Keim et al., 2008). Visual analytics combines a number of areas (Thomas and Cook, 2005), including:

- Data representation and transformation – Converting heterogeneous and dynamic datasets into a usable format.
- Visual representations and interaction – Using information visualization and visual perception to explore data.
- Analytic reasoning – Supporting “assessment, planning, and decision making” through analysis techniques and algorithms.

- Production, presentation and dissemination – Communicating analysis results to a variety of audiences.

This definition for visual analytics details a holistic framework for analyzing data starting from the moment data is gathered through to the final disseminated results gain from analysis. In contrast to the visualization domains I covered in chapter three, visual analytics intentionally brings in accessory actions - data gathering and dissemination - in collaboration with visualization and analysis. Other visualization domains may, technically, already perform these accessory actions but, as the proponents of visual analytics argue (Thomas and Cook, 2005), describing a holistic approach is not an overall goal of these visualization domains (although, critical visualization may be an outlier in this case because it seeks a holistic approach to visualization).

Visual analytics is therefore seen as a comprehensive empirical/analytical approach to interpreting and analyzing data using visual representation. However, I am concerned Thomas and Cook's definition of visual analytics does not describe the extent of how 'visual analytics' can be applied. As I laid out in chapter three, many of the visualization domains described have their own representation methods and attributes. Some visualization domains question how visual representation should be handled and may not conform to the standards laid out by the information visualization domain that is cited in Thomas and Cook's definition.

Analytic reasoning as a data analysis method is also questioned by domains like artistic, casual, personal and playful visualization. Many of these visualization domains allow for other data analysis methods such as creation, monitoring, reflection and reminiscing. In an attempt to combine these contrasting viewpoints I must alter the definition of visual analytics with an intention of incorporating the differences between the visualization domains. While relying on information visualization and analytic reasoning is a valid approach to data analysis there should be other avenues we can take when using visual analytics. I define visual analytics as:

A practice where methods and systems are developed to gather, visually represent, interpret and disseminate data for analytical, creative, monitoring, reflective and/or reminiscing purposes.

This definition keeps Thomas and Cook's holistic framework but makes sure other means of interpreting and experiencing data are covered.

With a new definition of visual analytics I can now begin to apply the practice to a specific domain. In this case I refer to the domain of game analytics, which I have in the past described as “the systems and methods used to analyze game-related data” (Medler, John, and Lane, 2011). I should note that game analytics does not need to use visual representation to analyze data but visualization is a major part of both games and data analysis. Not only do games make use of visualization as part of the process of playing, visualization can also be useful for analyzing game-related data, as I have already argued. Thus, I seek to understand how the practice of visual analytics has been and can be applied to games as part of game analytics.

Using my definition of visual analytics above, game analytics can be defined as being useful for many data analysis purposes (analytical, creative, monitoring, reflective and/or reminiscing). However, game analytics carry with it the biased that only analytic reasoning is the valid method of data analysis (in a similar way visualization domains, like information visualization, have a bias towards analytic reasoning). Game analytics should be used to study game-related data and find insights into how games can be altered, if it were to only focus on analytical data analysis. In this respect, a game's developer benefits the most from game analytics. Developers can use game analytics to analyze player data to understand what features should be added to their game or analyze data from their game's source code to determine how to make the code more efficient. Other data analysis methods, such as reminiscing, are not as important to game developers when compared to the analytical data analysis method because they would not produce ‘actionable insights’ that have immediate impact on a game's development. However, those other data analysis methods may be important to other audiences.

Players, for example, may wish to reminisce about their past gameplay or reflect on the strategies they have been using in a game. Spectators may wish to monitor their favorite players or teams. Game analytics should not be limited to analytic reasoning alone.

In order to present game analytics as both a practice where analytical data analysis and other forms of data analysis are used I split game analytics into two domains. The two game analytic domains I discuss are development analytics – which has an audience of game developers and is routinely more analytical - and play analytics – which has an audience of players, spectators and third-party service and use a mix of data analysis methods. Development analytics refers to the methods and systems used to interpret data by a game's developer, which includes every profession or audience who helps release/maintain a game: marketing, financial, programming, design, customer service, etc. The methods and systems used by the developer audiences are usually kept internal to the group or company where the game is being developed. Furthermore, game studies research work that analyzes game-related data routinely falls under the heading of development analytics too. Research often focuses on understanding players or a game's system in order to improve the development process; hence this type of research aligns with the motivations of game developers. I should note that development analytics is only discussed in this chapter.

Play analytics is the main domain I explore for the rest of this dissertation. As opposed to development analytics, play analytics refers to the methods and systems used to interpret data by those who play or spectate a game. Play analytics is therefore important to outside parties: players, third-part services and spectators. These outside parties seek to analyze data from games for their own purposes, whether those purposes relate to playing a game (players), providing services for players (third-parties) or watching players (spectators). The key difference separating development analytics and play analytics is whether the audience using an analytic system is interpreting data with or without a motivation to alter the core development of a game itself. Development analytics are used to alter a game, play analytics alter the experience of a game. Play

analytics focuses on motivations like reflecting, sharing, competing, monitoring and creating, all coalescing around the motivation to play and experience games.

In this chapter I present examples of both development analytic and play analytic systems. The development analytic examples are meant to show the type of systems currently being used by developers and researchers in regards to interpreting game-related data. Understanding the types of tools developers use gives me a baseline to refer to when discussing play analytic examples. This baseline helps discuss how play analytic systems are built to allow for different data interpretations and experiences in respect to systems made for developers.

This chapter begins with examples of development analytic systems. General analytic tools available off the shelf are discussed in regards to business-related analytics. There are also example ad hoc analytic systems developers have built for the purposes of iterative design, while other systems have been built to operate within a game's engine and provide better software debugging support. These examples are not meant to be exhaustive but to give an overview so I may refer to these systems throughout the rest of the dissertation. I also list a set of properties defining development analytic systems in order to contrast the systems with play analytics. Next, I describe a number of keystone play analytic systems which I view as being important to the domain. Developer produced and player created systems are covered, each looking at different types of data and promoting different analysis methods. A list of properties describing play analytics is provided too and compared with the properties of development analytics. Finally, the end of the chapter is where I discuss how play analytics can be interpreted by the functional/phenomenological framework I have used to interpret data and visualization. This means looking at how the functional properties – visible, comparable and control – affect play analytic system design and how the phenomenological properties – exposure, interconnection and historic – can be used to interpret the experience players have when using play analytic systems.

Development Analytics

Game analytics has many domains and audiences much like the visualization domains I discussed in chapter three. There are numerous game developer audiences for whom game analytics can benefit. For game developers, everyone from the financial analyst to the game designer can make use of game analytics in some way. Data detailing sales reports, forum posts, player behavior and source code are each important for their respective developer audiences. These audiences are looking to analyze game-related data in order to make important decisions on how to make a game more efficient, more profitable or, otherwise, better. Conversely, players do not necessarily wish to use game analytics to make a game better, hence my decision to separate the discussion based on these two audiences (developers and players). Players instead may wish to use data to become better players, create art, show off their achievements or tell stories. Regardless, both audiences, developers and players, have a diverse spectrum of available game analytic tools at their disposal.

Below I break up the tools that developers use into three areas related to business, design and performance. This is not meant to provide an exhaustive list of tools used for development analytics but meant to show three areas where development analytic tools are used. First, in the business and marketing section I look at the tools developers can buy off the shelf. These tools are typically designed with generality in mind and provide standard templates for data analysis. Business and marketing analytics tend to follow standard methodology when analyzing games thus tools that provide a consistent experience are helpful. Second, other game developer professions choose to build their own analytic tools in order to analyze their design process. Programmers, designers, artists and producers may require features that other general analytics tools do not provide. Hence they need to build their own ad hoc tools in order to modify the feature sets available. Third, not only are some analytic tools built by game developers but those tools can be built into a game itself. Building analytics tools to work concurrently to a game at run-time can help developers, like AI programmers, understand how their game is performing in real time.

Business analytics

Many companies, regardless of being in the games industry, use business analytics. In his book *The Numerati* (2009), Stephen Baker chronicles a number of instances where data analysis is being used for everything from supply-chain management to medical monitoring. In each of these areas standard methods and analysis tools are being developed and deployed to aid the data analysis process. Game development is no different. Game analytics has been gaining momentum over the years and the availability of general analytic tools for analyzing data regarding business development and marketing research has been growing too.

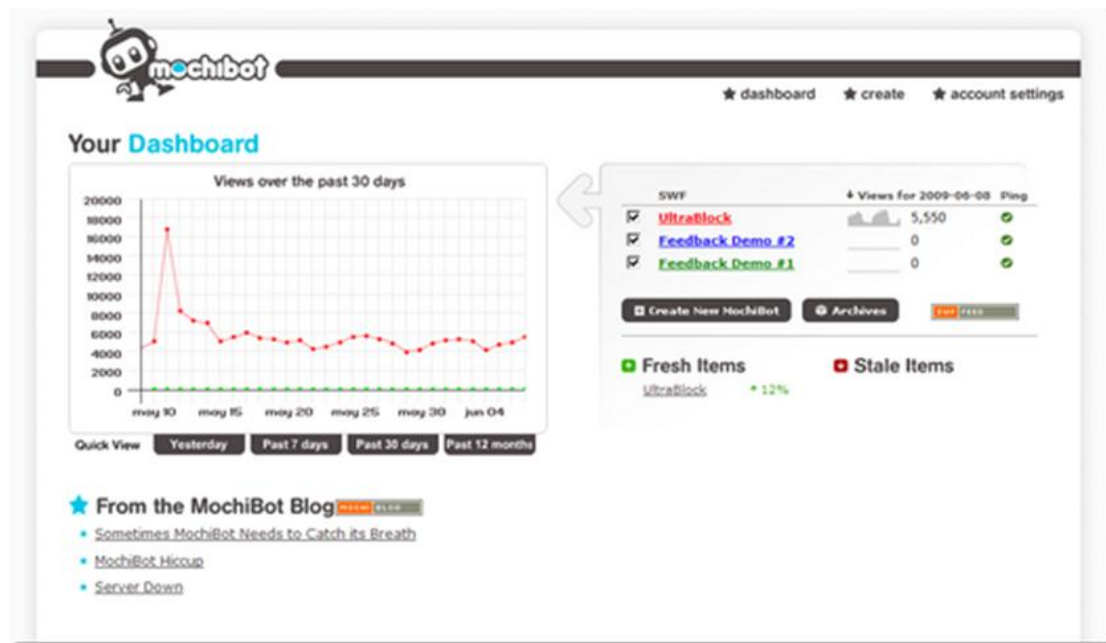


Figure 4.1: Mochibot (pictured) is an game analytic system used by game developers to track games built on the Flash platform.

Perhaps the first real push for widely accessible business analytics for games was on the Adobe Flash platform. For many years, online services such as Mochibot (Figure 3.1), Nonoba and Playtomic have provided Flash game developers with tools to track game data. In the beginning, Mochibot, one of the first companies to offer developers a simple service, tracking how many players loaded a game in order for developers to show advertisers how many “game views” they had over a period of time. The service at that point was about ad revenue more so than gameplay tracking. As time passed, and other Flash tracking services began emerging, these services began offering more ways for

developers to use data for understanding how to better monetize and design their games. Tracking now includes measuring when users click on items, say a specific ad, or how long it took players to complete a level (including a number of other possible gameplay measurements). Dashboards (webpages built to display charts and graphs) and visualizations offered on each service's website display the collected data over time, giving developers the ability to monitor their performance or test their design ideas. Many of these tracking services are also available for free as a starting package, something that even novice or independent game developers can afford. This makes it easy for developers to begin using the services without any real risk. Unfortunately, for developers on other platforms at the time these services only existed on the Flash platform.

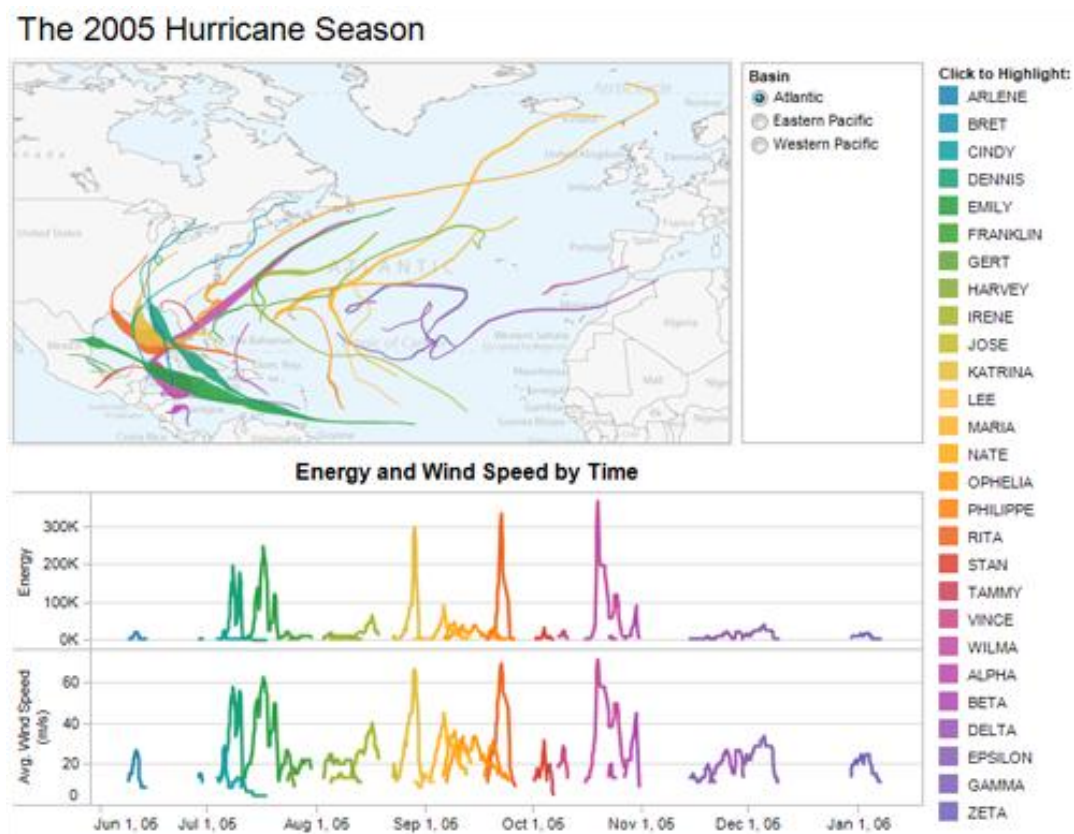


Figure 4.2: Tableau (pictured) is a general purpose analytic tool game developers can buy off the shelf.

Non-Flash developers, having no services providing general analytic tools, had to build or buy their own analytic tools if they wished to track their game. Over the years this has changed. Today there are a number of general analytic tools available for game

developers to use as part of their business and marketing analytic strategy. Tools such as Tableau (Figure 4.2), Spotfire and Mircostrategy are built to be general purpose visualization tools that can be connected to multiple sources of data and servers. Considering that business and marketing analytics typically use the same methodology for each game (tracking sales reports, player retention, player affinity, etc.) using these tools make it easier to create a common set of dashboards or visualization to use across any number of games. Even a program like Excel, a tool used ubiquitously in the games industry, includes features for visualizing spreadsheet data. All of these tools, like the services offered for games on the Flash platform, attempt to make analytics easier for developers to approach.

Besides business analytic tools being used to analyze player behavior in games there are social media analytic tools for analyzing how players are discussing games online. Online social networking has grown right along with game analytics both as a new medium for social communication and as an area to tap for investigation. Social media analytic companies like Radian6 and Looxii gather data from feeds being pushed out by various social media sources. Social networks, blogs and content sharing websites are monitored in order to find keywords related to brands or companies. From a marketing standpoint, having access to these feeds gives a fairly large encompassing window into what people are discussing and reacting to online. These social media analytic tools are made to be general too because any company wishing to capture data from social media feeds ultimately uses the same sources. For game marketing, a single social media analytic tool can be used to track any number of game titles or developers.

Iterative Design Analytics

Iterative design, a design methodology based on consistently prototyping and redesigning a product, is a widely used method in game design. Iterative design is also a design methodology that happens to work well with game analytics. Once a prototype is produced as part of the iterative design process, data can be collected regarding the evaluation and performance of the prototype. Collected data is then analyzed and decisions are made determining how to redesign the prototype, if need be. Iterative

design analytics therefore fall under development analytics as an important domain for developer audiences designing the game: producers, programmers, artists and designers. Technically, tools like Tableau, Spotfire and Mircstrategy can be used for analyzing data related to the iterative design process. However, having already discussed those tools I present here samples of ad hoc tools built from scratch by game developers to help iterate on their designs. What these tools show is how producing an analytic system from scratch allows a team of developers to include features other general analytic tools do not offer.

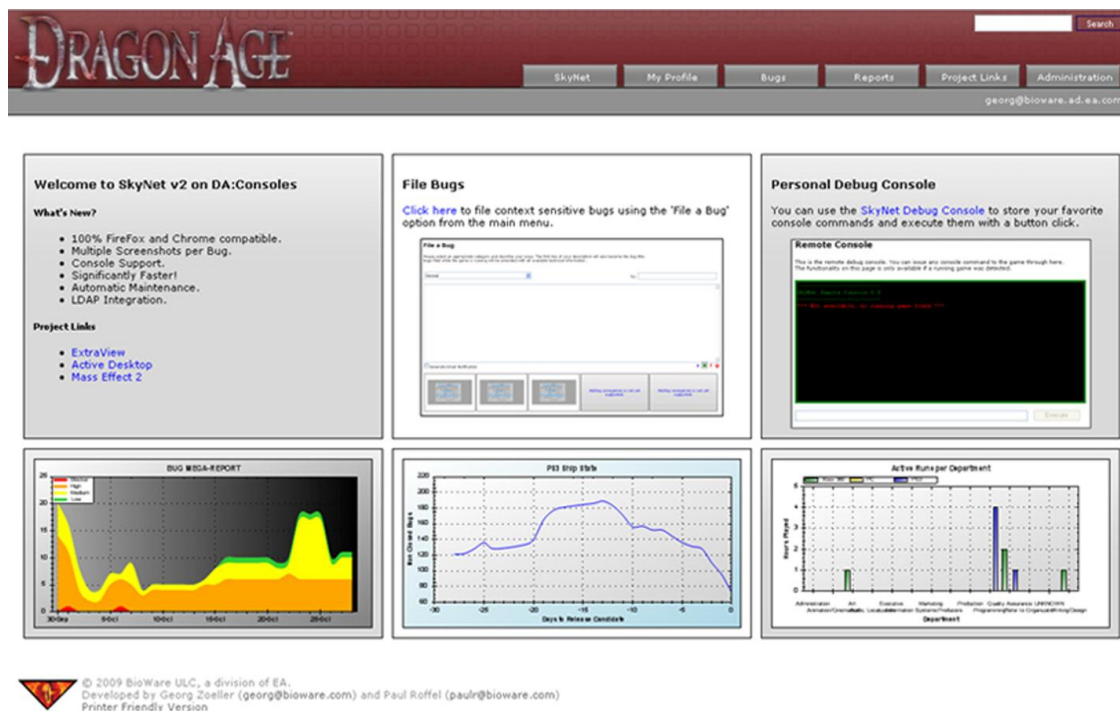


Figure 4.3: SkyNet (pictured) is an ad hoc game analytic system built by Zoeller (2010) at Bioware and is used to monitor developer behavior.

First, Microsoft's TRUE system (Kim et al. 2008) is an example of a visual game analytic tool built to aid user researchers as they play test games. When used in conjunction with play testing, the TRUE system records video feeds of each player testing the game, records game metric events and prompts the players with surveys to gauge their attitude during play. Video and attitudinal data contextualizes the captured game events providing user researchers with a greater understanding for why player testers were behaving a certain way in the game. The TRUE system can also be externally deployed during beta tests, continuing to capture game events but losing the

ability to capture video. Data that is captured after any test is visualized by the tool using various charts, graphs, heat maps, etc. If video was captured the tool synchronizes the event and attitudinal data with the video playback, making referencing all three data streams seamless. The tool's visualizations are then used to create analysis reports or allow game designers to drill down into the data themselves to find relevant information. Video and game event capturing are not features routinely available when using general, off the shelf analytic tools like Tableau or Microstrategy. Building the system from scratch allowed the developers of the TRUE system to incorporate those features into one cohesive system.

A different tool built to aid programmers instead of user researchers is the Skynet system developed by Zoeller (2010) at Bioware (Figure 4.3). This system aids the game development process by monitoring developer behavior rather than player behavior. Bug tracking, software metrics and social networking features are combined to create an online-portal where developers can both stay in touch with one another and have consistent access to their game's development status. While many features in SkyNet use spreadsheets to represent data, visual callouts such as color-coding important information are used, in addition to other visual graphs to display information such as the number of bugs filed and fixed. Heat maps are also used to denote where in a game's virtual space the most crashes occur, even giving developers the option to click on the map in SkyNet and be transported to the location within the game environment. Features like bug tracking, social networking and links directly to the game environment are rarely available in other general analytic tools.

Finally, other visual game analytic tools focus on the exploration aspect of play and attempt to provide an attractive environment for data analysts. Data Cracker (Medler, John, and Lane, 2011) is one such example, which I personally helped build for Visceral Games to analyze player data from the game Dead Space 2, a horror themed first-person shooter (Figure 4.4). The tool was built to be aesthetically pleasing, accessible to the entire game team and yet still have some semblance to other general analytics tools such as Tableau. For example, color schemes and artwork from Dead Space 2 are used to

brand the tool to the team which can increase the inherent attraction of the tool as well as information retention. The tool was also made more accessible by providing visualizations in a tiered format which becomes more detailed as a user drills down into the data. Graphs consisting of, for instance, the types of weapons players use in Dead Space 2 are displayed in tiers that step analysts through a general overview of the player population down to how weapons are used on specific maps or by specific player groups. This allows team members to become integrated with the tool over time instead of presenting them with an overly complex visualization when they begin. Altering the aesthetics of a tool and making sure the tool is accessible helps motivate team members to get involve with the analysis process and hopefully allows more people to play with the tool. While general analytic tools like Tableau do provide methods for altering the appears of the visualization templates it provides, having complete control over the aesthetic look of a tool is very different from altering a few display criteria found in most general analytic tools.

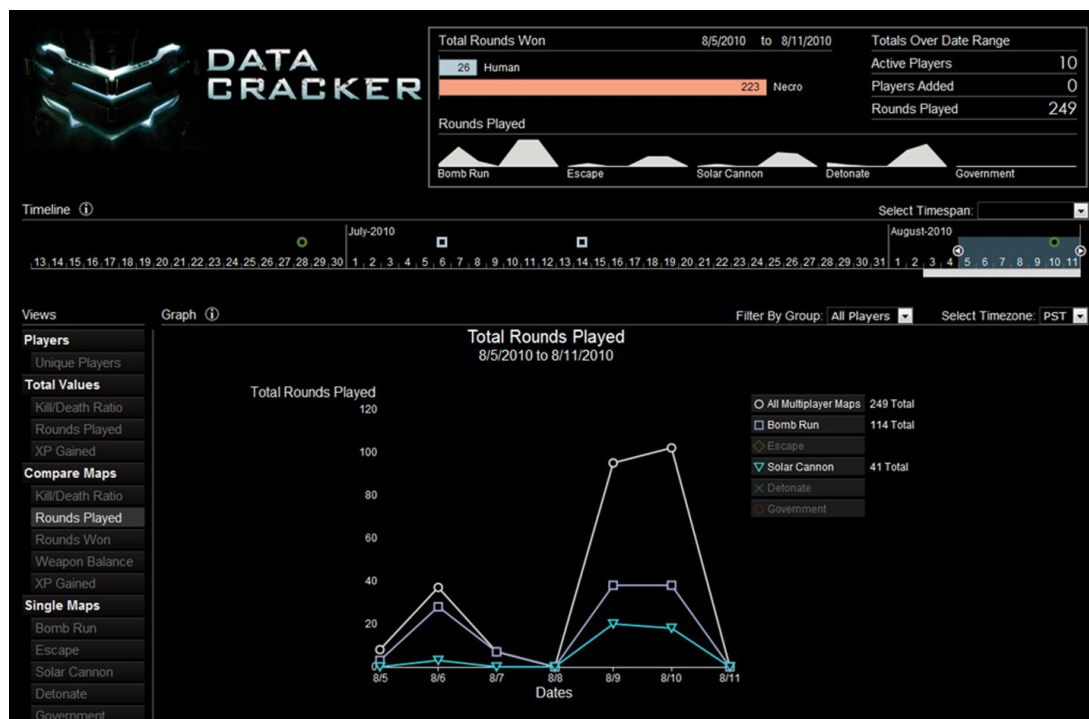


Figure 4.4: Data Cracker (pictured) is an ad hoc tool developed to monitor player data from the game Dead Space 2.

AI Debugging Analytics

Debugging software involves searching for failures or other misbehaving code within a program. It is a process that tends to consume most of a programmer's development time (Beizer, 1990). Sometimes debugging methods like break points are used to stop a program while it is running and display data related to the current state of the program. Game programmers, for example, may set a break point to stop their game when a player or non-player character (NPC) performs an action because the action is causing an error. When the program stops running they can review the game's current state data to determine what is going on. However, a different method is to provide data detailing the game's current state at run time and to make debugging part of actually running the game.

Real time visual analytic tools built into game environments can be used for debugging software. Examples of these tools have been built for games such as *Kingdoms of Amalur: Reckoning* and *The Sims: Medieval* (Dawe et al. 2011). AI programmers Dawe and Graham argue building real time debugging tools for visualizing AI behaviors benefit the development process greatly. First, valuable real time game data is visualized using their tools. Data associated with navigation, character pathfinding, line-of-sight, character attributes and available object interactions can be visualized within the game environment. Second, visualizing data in real time allows debugging in real time. Any data visualized provides an immediate indicator of whether or not the data is correct. If a character's visualized line-of-sight points to the ceiling it becomes apparent why a character cannot see an object directly in front of them. Real time debugging also gives programmers the freedom to debug a problem on any machine. If a visualization tool is built into the game environment and a bug appears on someone's machine, a programmer can debug the issue on that machine instead of attempting to recreate the problem on their own development machine. This leads to the third benefit, a built-in visualization tool means the tool can be used by anyone. Designers adding character attributes in the game can monitor how those changes affect the game without having to wait for someone else to test the changes, for example. Finally, visualizing real time data can be paired with recording data. The tool built for *The Sims: Medieval* both

visualized data and recorded data related to AI behaviors like pathfinding (Figure 4.5). Once a NPC falls off their correct path there is little need to analyze their current state beyond acknowledging something has gone wrong. Recording data, having a history of data over time, allows a programmer to step through a character's pathfinding steps in order to find the moment when their pathfinding went wrong. Visualization provides an immediate acknowledgement of the game's current state and recording data over times means developers can step back to review how the current state came to be.



Figure 4.5: A screenshot of the visual AI debugging and analytic tool built for The Sims: Medieval.

The navigation mesh, which describes where in-game characters can walk, is shown.

Properties of Development Analytics

General, ad-hoc and concurrent are three terms that define the different development analytic tools I covered. General tools are available for conducting game analytics and allow for a uniform experience across games. Some developers build their own ad hoc tools instead, opting to maintain their own analytic tools. Finally, concurrent tools are built directly into a game's engine to provide real-time analysis as part of the development process. These three terms are useful to refer to as I discuss the properties that define the domain of development analytics.

The following development analytic properties revolve around the fact that development analytic tools help developers analyze their game as part of the development process. Any game being developed can benefit from data analysis because developers must consistently evaluate their games. Data analysis tools provide developers with a way to make informed decisions on how to alter their games based on the data they collect about their players and game content. I step through the properties defining the domain of development analytics below and I use these properties as counter-points in the next section when I describe the set of properties for play analytics.

- 1) Development analytic tools are built for developers or third-parties concerned with the development and the continued maintenance of a game.*

Development analytic systems concentrate on aiding the development or maintenance of a game. The systems can be built internally, in which case a system itself may not be as polished as a system meant to be used by a larger audience. Skynet for example was an ad hoc system maintained by the developers using Skynet at the time and therefore only met their needs, instead of the needs of outside audiences such as players. Other times general analytic tools are acquired and used for game data analysis. These tools often include the ability to share dashboards or visualizations created using the tool, which is meant to facilitate collaboration among multiple developer audiences. All of these tools are development focused and used as part of the development process.

- 2) Data used for development analytics is typically gathered directly by developers and is not necessarily made accessible to outside parties.*

The audiences for development analytic systems, i.e. the developers, typically have control over the entire analytic process, from gathering data through dissemination. For example, an AI debugging analytic system is typically built by the actual AI programmers using the system. Those AI programmers control what data is gathered,

how it is visual represented and use the tool for analysis. Other times a development audience may have to rely on another development audience to create the methods for capturing data. Designers, for instance, may not necessarily build the programs used to capture the data they need but, since they are part of a development team, have access to other developers who can build the data capturing programs, say software engineers. The Data Cracker system described above had one team developing the data capturing system while another team developed the analytic system. However, the analytics team constantly conversed with the data capturing team and requested changes to the data capturing methods over time. Access to the data capturing method and having the ability to change how or what data is collected differentiates development analytics from play analytics.

3) All forms of game-related data are used for development analytics.

Different development audiences may wish to use different forms of game-related data, not just data related to gameplay. Flash designers may use a service similar to Mochibot to track how often their game is played, game programmers may use a system like Skynet to track bugs or a public relation team may track social media feeds to monitor the brand awareness of a game. Data used in development analytics does not need to directly relate to actual player gameplay but covers any type of data related to the process of game development.

4) Development analytic systems typically follow the methods used in the information, scientific and minimalist visualization domains.

The systems described above use both information and scientific visualizations to aid in data analysis. Systems like Data Cracker or the general analytic tools (Tableau, Spotfire, etc.) use information visualization to visualize abstract data with no tangible visual representation like weapon damage or the number of times a level is completed. Other systems like Skynet or the Sims AI debugging tool use scientific visualization because they use data related to game spaces with a visual representation, such as a map

or level. Minimalist visualization techniques are typically paired with either information or scientific visualization because the data visualized should not be distorted in anyway. The data should provide value for the developer who uses it to make decisions affecting the development process. While this does not mean other visualization domains could be used, information, scientific and minimalist match the analytical data interpretation style of development analytics.

5) Data is often interpreted analytically or is monitored as part of development analytics.

Development analytic tools aid in the development process by providing analytical interpretations of data that produce insights into the development process. Common tools like Mochibot, Nonoba and Playonomics provide a standard set of analytical interpretations for Flash developers such as how long players are spending in their game. Skynet allows developers to monitor bug reports and links them to a player's position in the game. Visual AI debugging tools allow programmers to view AI script behaviors in real time in order to see how the AI agents are sensing the game world. These are not tools that are casual or playful, they are work related tools built for audiences to actively use over long periods of time and to aid in the process of sensemaking.

Play Analytics

Play analytics is a domain that seeks to understand how players, rather than game developers, use data analysis as a method of play. Play analytic systems are built to promote or support play through a combination of data collection, visual representation, analysis, interaction design, user experience and product dissemination. These systems are more diverse in the kind of analytic experiences they offer compared to developer analytics. Where developer analytic tools are meant to interpret data analytically (or monitor data), play analytic tools additionally allows for player to create, reflect, reminisce. Below I cover a number of play analytic system examples to provide an initial view of the type of tools that exist in the domain. Each tool has a different approach to

play analytics and these differences help convey the diversity of play analytics. I end this section with a list of properties that define play analytics. These properties are also compared with the previous list of properties defining development analytics.

Bungie.net

Built for presenting data captured from Bungie's Halo game franchise, Bungie.net offers a similar level of meticulous player dossier record keeping compared to other professional sports (Medler 2009). Every kill made, point earned and objective completed by a player is recorded during a battle, from both single player and multiplayer gameplay. Player data is visualized using a number of visual forms including: line graphs for point progression over time, heat maps for detailing where events occurred in relation to the game map and sunburst graphs detailing when events occurred during separate sections of particular game. These visualizations are only offered within the online system and most player data cannot be accessed through the game. Limiting data access to Bungie.net creates a type of trophy room interface to each player's data. Awards and gameplay are put on display to explore outside of the game and far surpasses the investigative potential of systems relying solely on cumulative score keeping or leaderboards.

Giant Bomb

Giant Bomb offers a third party play analytic system built to aggregate both game information content and player achievement data (Gerstmann and Davis 2008). A game news, wiki and community website, Giant Bomb generally relies on their community of game players to add to the site's massive collection of game information content. Players can add content about a game's mechanics, stories, characters, developers, etc. In this respect Giant Bomb actually acts as a play analytic system that collects data about games. Although, one specific feature of the website is related to aggregating data from different achievement systems that exist on a number of gaming platforms (Figure 4.6).

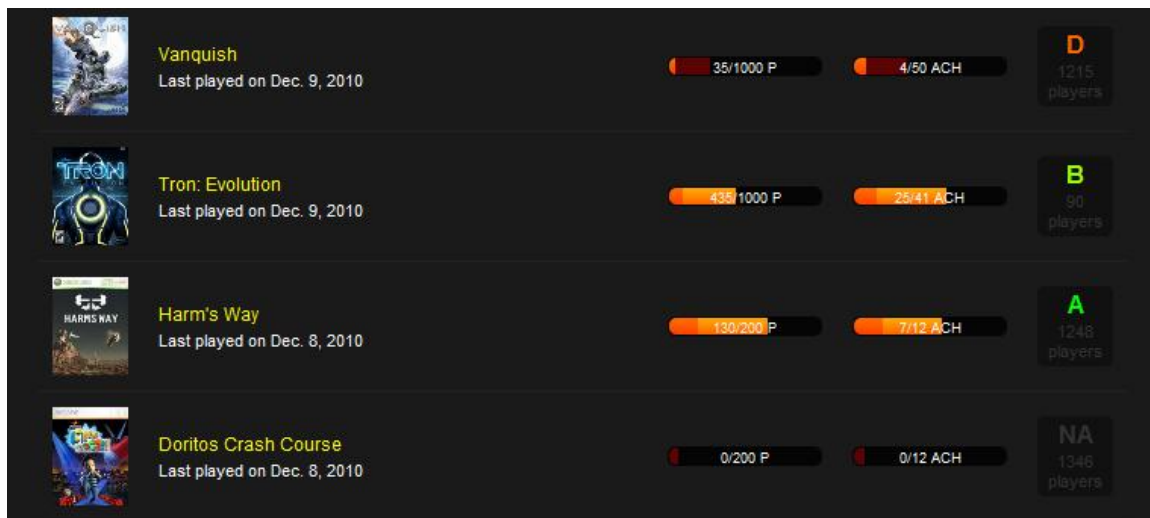


Figure 4.6: Giant Bomb aggregates player achievement data from multiple platforms and uses it to, among other things, find the average achievements earned per game.

The achievement aggregation service Giant Bomb offers is a way to bring a player's achievement data from different platforms into one cohesive system. As a member of the Giant Bomb community a player can connect their achievement data from four different sources: Playstation Network (Sony), Steam (Valve Corp. 2003), Xbox Live (Microsoft) or World of Warcraft. Achievements from these sources are combined and visualized allowing a member to view all of their achievements in one place. Access to each member's achievements also gives Giant Bomb the power to determine additional information about achievements that would normally not be possible. For example, the rarest achievements can be determined which can add additional value to those achievements. The average number of achievements earned per game can be found too, giving another piece of data that can be visualized to members for comparison with their personal achievement data. Combining different achievement datasets essentially allows Giant Bomb to create "a visual achievement catalog for their members to peruse achievements from a different perspective" (Medler and Magerko, 2011).

Sims Exchange

The Sims Exchange (Electronic Arts Inc. 2004) is one play analytic example where players can use past gameplay data to create their own stories related to the sims families they create in the game The Sims 3. The Sims Exchange is a service for sharing assets built for games in the sims franchise. Additionally, players can create stories about

their sims characters and share them on the exchange (Figure 4.7). These stories are created in a slideshow format where screenshots from the game are combined with text detailing the narrative. No player data is recorded automatically, making The Sims Exchange different from other play analytic systems like Bungie.net. Instead, players can dictate their own play experience with their own personal interpretations of the game events overlaid on top. It's a qualitative form of data analysis. Considering The Sims 3 is supposed to represent playing with a doll house it makes sense that the players are given the ability to tell their playful stories using The Sims Exchange.

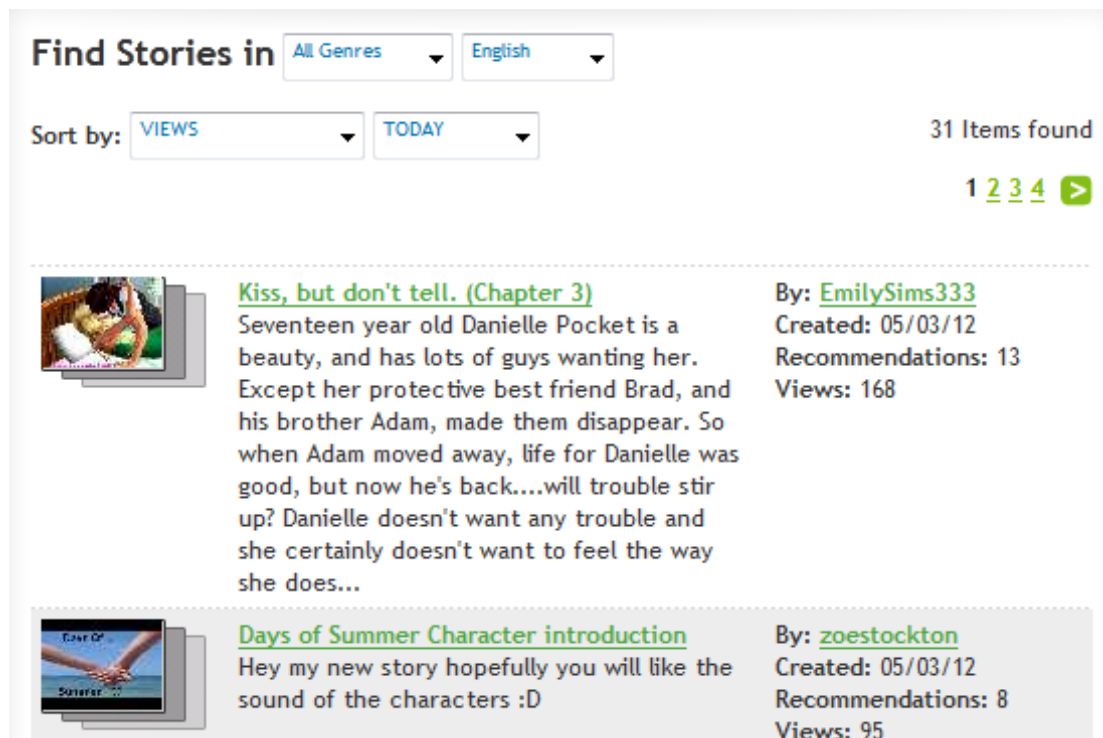


Figure 4.7: The Sims Exchange allows players to create their own stories using screenshots taken from The Sims 3.

WoWhead

Some players band together to create large databases filled with data about specific games when no other similar online resources exist. Websites like WoWhead (WoWhead 2006), which specifically collects data based on the game World of Warcraft, contain massive amounts of data regard in-game content like monsters, items, quests, non-player characters, etc. This data is often visualized in relation to maps pulled from the game to help players find a specific monster, item or quest. The World of Warcraft

website does not provide such detailed data about their game which is why players have created this style of game database to function as an optimized way for players to share information about the game and help players that may be stuck or wish to find a specific piece of game content.

SC2Gears

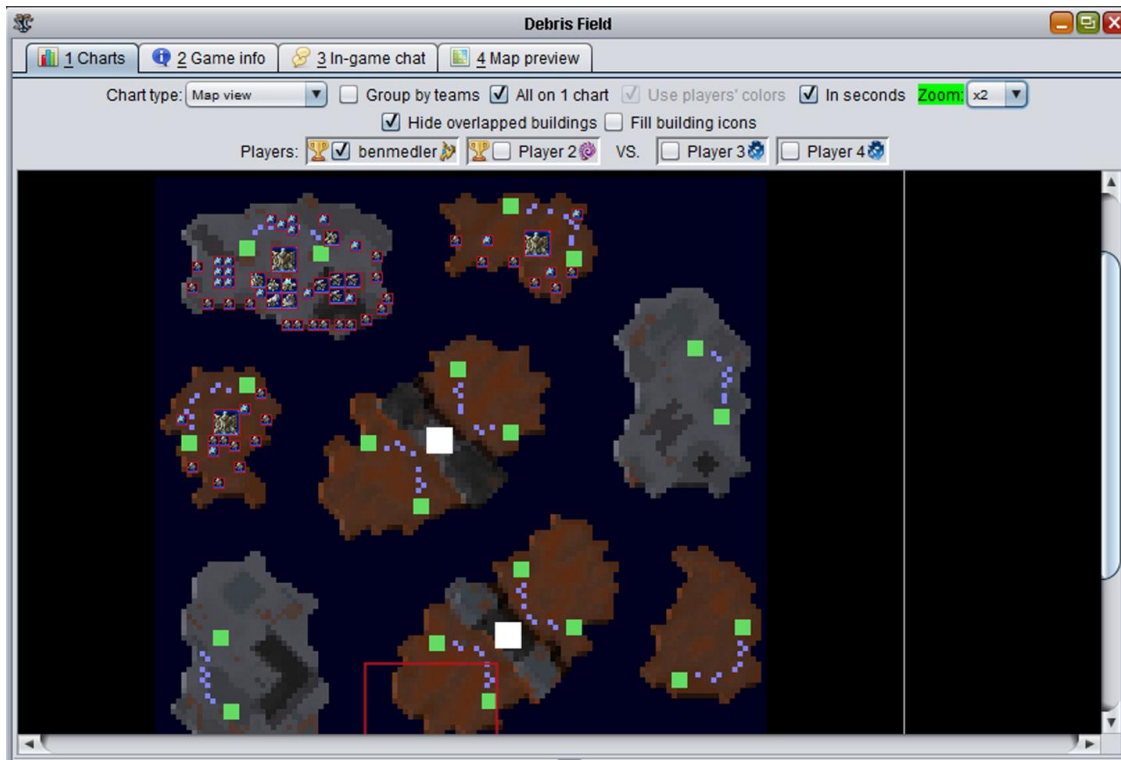


Figure 4.8: SC2Gears is a player built ‘replay analyzer’ that parses Starcraft 2 replay files and provides a series of visualization and analysis reports covering the replay data found in a file.

Different games allow players to record the events of a play session and save those events as a replay. The replay can be viewed after the play session concludes and works like a recording of the session. Players can fast-forward, rewind and review the events that occurred. Replays contain a wealth of data and some players have built play analytic tools to analyze replay files, like those used in the Starcraft game series. Starcraft 2 is a real-time strategy game where players build manufacturing facilities to produce military units and battle one another across an in-game map. Replay files from Starcraft 2 record player actions such as building units, attacking other players or moving their field

of view (i.e. the player's camera). Expert players, or players wishing to get better at the game, revisit their replays in order to analyze their own strategy (much like sports teams analyze video tapes of their games). Normally, replay files are reliant on the game's engine to reproduce the data contained within the files and only a single replay can be played in the game at once. However, Sc2gears (Belicza, 2010), a replay analysis tool (Figure 4.8), takes those replay files and visualizes the information outside of the game environment. The tool allows players to analyze multiple replays at once, which is not available in Starcraft 2. Players can view a 2D graphic visualization of where they built their facilities on the map, monitor their 'actions per minute' count, and review the order they built their units and facilities. SC2Gears help players reflect on their past gameplay and compare multiple replays at once (a feature unavailable to players until SC2Gears was built).

AutoLog

Some play analytic systems are actually built right into a game and compare a player's data with their friends' data. Criterion's Autolog system (Electronic Arts Inc. 2010), built within the racing game Need for Speed: Hot Pursuit, is a recommendation analytic system for promoting competition. Autolog keeps track of a player's data related to their races (generally this means their fastest race time) and compares their racing data with their friend's data. These comparisons are automatically turned into challenge recommendations, which are visually displayed within the game environment. Players are asked to beat the score or time of one of their friends receiving a reward for completing the recommendation. This creates an improvised playful environment amongst a community of friends where players choose which goals and competitions to pursue.

Properties of play analytics

Similar to development analytic tools, there are general, ad-hoc and concurrent play analytic tools too. The difference, however, is the line between what is considered a general, ad-hoc or concurrent tool begins to merge in the domain of play analytics. A general play analytic tool, used by a diverse audience for diverse reasons, can be built by developers (Bungie.net) or can be built by players (WoWhead). Tools that begin as ad-

hoc tools like WoWhead and SC2Gears, which are built by players, can turn into general tools for other players to use. Finally, developers have an easier time of building concurrent tools, like Autolog, compared to players since the developers have access to the game's source code. However, players can build tools like WoWhead which contains valuable information about a game, in this case World of Warcraft, and other players can take advantage of the information as they play the game in real time.

I argue the reason why we begin to see this blurring effect between general, ad-hoc and concurrent play analytic tools is because play analytics is a diverse domain covering many of the data interpretation methods (analytical, reflection, reminiscent, etc.). These tools are built for players, all of whom have different motivations for playing and using play analytic tools. But these tools also focus mainly on gameplay or content found in the game. When compared to development analytic tools, which can be used for many diverse game-related data sets, play analytic tools seem to be of limited scope. Although, perhaps it is the fact that play analytic tools revolve around gameplay that makes it easier for them to blur between being a general, ad-hoc or concurrent tool. Every player ultimately wants to play, these tools help facilitate that motivation.

Below I list the properties defining play analytics. While these properties concede that play analytic tools cover a diverse set of data analysis methods, these tools remain subject to control over what data is available from a game. This control is either put in place intentionally by game developers, especially when those developers provide a play analytic tool, or can be the limitation players have in regards to collecting data from a game when they do not have access to the game's source code. Some tools do find ways to break away from any restrictions a developer may erect in order to control what data is made available. SC2Gears is able to deconstruct Starcraft 2 replays, for example, and in some cases a developer willingly provides access to raw data, such as Bungie.net. Regardless, to be considered a play analytic system some form of analytics must be provided to players, whether through a game's developer or not.

- 1) *Play analytic systems are built for players, spectators or other non-developers.*

Play analytics is a domain for players or the audiences who do not affect the overall development or maintenance of a game. Dissemination thus becomes a bigger issue than with developer audiences. As we saw examples in the development analytics section the tools developers use can be constantly rebuilt or restructured in an ad-hoc manner to match the needs of the developer. Data can be linked to general tools like Tableau for quick analysis or systems like Skynet can be created as a grassroots development tool where the developers maintain the system themselves. Conversely, play analytic systems are products non-developer audiences must use. These audiences may not be given, or have, the capabilities to modify a system or capture different data. A level of polish is thus required when disseminating a play analytic system in order to ensure the audiences can gain some sort of meaningful experience from the system.

- 2) *Play analytic audiences must have access to data, be given visualizations and/or provided with analysis tools.*

Beyond just polishing a play analytic system, audiences must have some level of access to the analytic process whether that be access to data, visualizations or analysis tools. If developers do not wish to build their own play analytic systems for players they can provide data feeds for players to build their own. If a developer chooses to build a system some visual representation of game data must be provided for an audience to use, even if the visualizations are static or otherwise non-interactive. Furthermore, systems may also provide an audience with analysis capabilities for finding patterns or correlations in the data. Any one or all three of these features must be provided within a play analytic system.

- 3) *Play analytics make gameplay data or other game-related data meaningful for audiences who play or spectate games.*

Marketing or game engine performance data is meaningful to development audiences but for a player these datasets are not what they wish to explore as part of their game experience. Gameplay data is typically the type of game-related data found in play analytic systems. Every example given of play analytics thus far has dealt with gameplay data such as scores, achievements, replays and play-through stories. Game reviews or data associated with a player's social gaming groups (e.g. forums or guild information) can be meaningful as well. It is hard to find examples of play analytics systems using marketing, server performance or financial datasets and even harder to find reasons why audiences like players would want to see these datasets (unless they had intentions to develop games).

4) Play analytic systems use the casual, personal or playful visualization domains, often being categorized as interactive, supportive or creative.

The key attributes of play analytics are similar to the attributes used to describe casual, personal and playful visualizations. Play analytic systems are almost exclusively for non-work purposes and portray personal player data. Within any system there can be a diverse audience consisting of players and spectators, given that a game can have a large, diverse player population. Play analytic systems are also built to be used for varying amounts of time. Some provide features to quickly reference game statistics or gameplay information, while other systems allow for long periods of data analysis. One example is SC2Gears, which provides Starcraft 2 players the ability to analyze their replays. Players can spend longer periods of time analyzing their data especially when compared to other systems like Autolog which are meant to provide players with quick recommendations to keep them playing. SC2Gears is also an example representing the three types of playful visualization. It is an interactive analytics tool, supports the analysis of Starcraft 2 strategies and is a project created by players hoping to analyze their data. Play analytics strives for diversity, providing diverse methods for analyzing data for diverse audiences.

- 5) *Analytical or monitoring data analysis methods are used but play analytics allows for reflective, reminiscent or creative methods too.*

While it can at times be difficult to define firm lines between the data analysis methods I have described (analytical, reflective, monitoring, reminiscent, and creative) it can be argued that there is a difference between which of these methods development analytics and play analytics tend to use. Development analytics seek to answer business and development questions when collecting data and often lean towards interpreting the data analytically. Play analytics is about providing players with entertaining or personally meaningful experiences and does not need to focus on interpreting data analytically. Earning achievements can be a meaningful experience for players but it does not require the players to interpret achievement data through an analytical process. This is not to say that an analytical interpretation cannot be useful. Giant Bomb's achievement system does provide an analytical interpretation of player achievements when they compare each player's achievements earned in a game to the community at large. At the same time, though, Giant Bomb's system allows players to aggregate their achievements from multiple platforms in one location, allowing players to reflect on their achievements from a holistic perspective. A perspective very few systems offer. The need to provide analytical insights through play analytics is ever present but is never the single ultimate goal.

The Functional and Phenomenological Aspects of Play Analytics

Play analytics represents a casual, personal and playful brand of analytics, in contrast to development analytics. Players and spectators are the audiences who use play analytic system and they do not have the capability to alter the core game features. They instead experience the game for themselves through gameplay. As a result, play analytics allows players to analyze, reflect and reminisce about their gameplay using systems that are made to augment their game experience. In order to discuss how players experience play analytic systems, this section provides an example of how to interpret play analytics using the functional and phenomenological framework I have been using throughout this first part of the dissertation. Below I describe how the functional properties – visibility,

comparability and control – manifest in play analytic systems. Additionally, I step through each of the phenomenological properties I presented in chapter one – exposure, interconnection and historic – describing how these phenomenological interpretations relate to the experiences players have using play analytic systems. Overall, I present how these functional or phenomenological properties can be used to describe, interpret and critique play analytic systems.

Functional Aspects of Play Analytics

Since both developers and players can build play analytic systems they must both worry about how the functional properties – visibility, comparability and control – manifest themselves within their systems.

Visibility

Deciding what data is visible should be the first task for anyone building a play analytic system because only visible data can be analyzed. What data is available can change from game to game but having an idea of what data should be made available makes it easier to decide how players can access that data. Bungie.net changed the amount of data given to Halo players over time, opting to show simple achievements and scores for earlier games (e.g. Halo 2) before creating detailed visualizations covering entire play sessions for later games in the franchise's history (e.g. Halo: Reach). WoWhead also evolved in a similar manner. Data related to World of Warcraft item, quest or NPC data was provided to players when WoWhead first started. These datasets offered a way to explore what content existed. As time passed and World of Warcraft players understood what content existed in the game, WoWhead changed to include more detailed information related to how to find and acquire content in the game. Locations of NPC and methods for obtaining a rare item were added to the site giving players a clear path towards accomplishing their goals. Additionally, deciding to hide data is another factor in play analytic design. The Autolog system only shows the race times of friends and hides scores from non-friended players. Designers believe friend scores are more important to players than the scores from strangers, hence the designers hide the scores of

other players. Therefore, designers need to decide what data is visible or exposed to players, what data is hidden and what data should be made visible over time.

Comparability

Comparability in play analytics becomes a question regarding what players or game content is being compared to other players or game content. Players are the typical audience for play analytics so they are the likely dataset to compare with game-related data. Are players compared against each other, against a sub-set of players, against computers or other constructed identity? Some games place players into a team, which can be compared as a whole, or a player's own gameplay can be compared against themselves. SC2Gears for instance allows a player to compare Starcraft 2 replays which may consist of their own gameplay data, gameplay data of their entire team or, if they didn't participate in a replay, may contain gameplay data from other players. Some systems automatically compare players together such as Autolog which generates friend leaderboards for comparing player scores. In other cases data representing game content is compared instead of players. WoWhead details thousands of in-game objects which players can sort and compare along a number of dimensions like damage, armor, magic attributes, etc. Furthermore, the stories players create on the Sims Exchange allow players to recommend a story to others but does not compare the stories directly with each other. Deciding how to represent comparability in a play analytic system means determining the ways to compare players or game content with one another.

The purpose for comparing data also needs to be established. Many games revolve around a type of competition or cooperation and comparing data within a play analytic system relate to these factors. A system can be setup to help players learn how to play the game and to become better competitors. Players could share strategy videos or comment on replay matches. Another purpose for comparing data allow players to reminisce about their own accomplishments such as Giant Bomb combining different achievement systems into one system for players to use. Designing how data is compared means creating the interconnections between players and/or game content while determining the reasons those interconnections are required.

Control

Finally, play analytic systems follow a process beginning with data gathering and ending with dissemination. Each step along that process has to be controlled by a developer or an audience. If a game developer is building a play analytic system then they may keep control over the entire process, as we see in examples like Autolog or Giant Bomb. These systems do not give players control over what data is used, how it is visualized, what analysis takes place or how the final data is disseminated. In other cases such as Bungie.net, players can access their data through an API. The API grants players the ability to take their data and create their own play analytic tool. Players may also decide to build their own analytic tools using game data they already have access to, like SC2Gears does with Starcraft 2 replays. There is also the question of how long data is kept and held by any party. If players can access their data, there may be a specific time frame when players can retrieve it before a developer must delete or remove access to the data. Having control within a play analytic system comes down to who has control over historic data records which are being collected from games over time and how those histories are being visualized, analyzed and disseminated.

Phenomenological Aspects of Play Analytics

The functional aspects of play analytics certainly have an effect on how a play analytic is used. However, players make a play analytic system their own, just like they make a game their own. Players have their own personal experience of the systems. We can interpret their experience with a play analytic system using the phenomenological properties – exposure, interconnection, and historic – and how the theories of bracketing, contextual interpretation, moods, historic context and prejudices contribute to those properties.

Exposure

Beginning with bracketing, designers of player analytic systems start with an ideal experience players should have within the system. Ideally, players should gain a sense of competition, for example in Autolog, or they should feel free to share stories as we see in the Sims Exchange. These ideal experiences should be reflected in what data is made visible to the players because it is likely a system's dataset will stay confined to the system itself. It may be hard to obtain or use data from systems like Autolog, Bungie.net, Sims Exchange and SC2Gears for other game-related purposes. I will not necessarily find it meaningful to combine the stories I create on the Sims Exchange with the analysis of my Starcraft 2 replays. I can therefore use bracketing as a method of interpreting a set of exposed data used within a play analytic system as the data meant to create the ideal experience for an audience. This interpretation method can also be used to discuss systems which aggregate data such as Giant Bomb. As a system that aggregates achievements from different systems I can interpret that the ideal experience for Giant Bomb is similar to the achievement systems the tool aggregates. However, since Giant Bomb aggregates achievements it can be interpreted that Giant Bomb bracket is larger than those of the aggregated systems and alters how a player experiences the system.

Altering the scope of a system's bracket (the amount of data it exposes) also changes how other audiences outside of the system may view the data. Games are not played in a vacuum, even though bracketing is a method for judging only the essential properties of an entity, disregarding other outside factors. Many play analytic systems expose data freely to the public at large. This may result in situations where I, personally, may not find any value in my game data outside of the context of the game but others may. There have already been incidents where player data from games has been used to find players who had committed crimes (Munsey, 2009), for example. Acknowledging what game-related data is exposed in connection to a player is something that needs to be covered when interpreting the types of experience players can have when using a play analytic system.

Interconnection

While bracketing looks at the ideal experience players should have within a play analytic system, contextual interpretation can instead be used to examine what affordances the available data in the system provides players and what connections are formed. Autolog leaderboards afford players the ability to brag or challenge each other in the game, to give a simple example. The Sims Exchange affords players the ability to tell a story about their gameplay or SC2Gears affords the ability to dissect gameplay strategies. Contextual interpretation provides a method of determining if the experience with a play analytic system matches the experience players receive from gameplay. SC2Gears and WoWhead are examples of systems created by players who felt there were no other play analytic systems which matched the experience of playing Starcraft 2 or World of Warcraft, respectively. In the case of SC2Gears, the replay system available in Starcraft 2 did not allow for robust analysis of a player's strategy (or the ability to analyze multiple replays simultaneously). Thus a disconnect occurred, the data players wanted was available but not contextualized in a way related to the experience of exploring strategies found in Starcraft 2. As part of my framework, contextual interpretation is used to interpret how the experiences players have in both game and analytic tool interconnect with each other, and whether there is a disconnect between those experiences.

Another method for interpreting interconnection within a play analytic systems is the mood of an analyst, i.e. players. As I covered in chapter three, there are many methods of data analysis available, including: analytical, creation, monitoring, reflection and reminiscing. These can be viewed as the type of mood an analyst can have when they are using an analytic system. For play analytics, this means players may approach a game or analytic tool with any of this data interpretation moods. For instance, on Bungie.net some players may wish to analyze their matches analytically, hoping to find better gameplay strategies, while other players may use the video capture feature in the game to share funny videos they made using one of the Halo games. Play analytic systems can also be designed to change a player's mood and alter their experience. Autolog is integrated into the gameplay of Need for Speed: Hot Pursuit, causing players to enter into a monitor mood where they constantly check their racing times and the racing times of

their friends. If the Autolog friend leaderboards were separated from the game it can be argued that the system would not make it as easy for players to switch over into a monitoring mood. In a similar way contextual interpretation is a method for discovering interconnections between game and analytic experiences, interpreting player mood explores how players connect and approach a play analytic system itself.

Historic

Gadamer's theory of historical context covers how players experience historic data within play analytic systems. Systems like Autolog, WoWhead or Giant Bomb are systems with very little historical data. Data is either aggregated and totaled over time (Autolog and Giant Bomb) or is presented as a static repository of game-related data (WoWhead). However, other play analytic systems do present data over time. SC2Gears shows a single Starcraft 2 match over time, giving players the ability to analyze the strategies used in that match. Bungie.net chronologically orders a player's matches giving players easy access to examine their recent matches. The real question, however, is why players need or would want to explore historic data related to their gameplay? As I examine more play analytic systems in the next chapter there are further examples of how placing game data within a historical context alters the experience players have in the system and in their gameplay.

The final phenomenological theory in my framework relating to the historic property consists of exploring the prejudices that exist within a play analytic system. Any play analytic system produced is built with some intention, whether they are the intentions of the developers trying to provide players with an additional system for experiencing a game or players building their own systems with the intention to augment their gameplay. I use Gadamer's theories about prejudices to critique the intentions behind building a play analytic system. Autolog, for example, was built with a prejudice that friend leaderboards are superior to general leaderboards. Players gain the advantage of competing with their friends but lose the chance to compete against a larger body of players. Building Autolog to focus on friends also create a situation where the Autolog

system becomes obsolete when a player with no friends. Building a system with specific prejudices can alter the experience of the system.

Prejudices can also be used to interpret player built play analytic systems. Players build their own system because a game's developer is not provide the system for the players. WoWhead is a shining example of a system providing players with data players desperately wanted but the developers of World of Warcraft did not provide. This is an example when the prejudices of developers conflict with the prejudices of the players. A developer wants players to explore their game, not find the fastest way to beat the game. Although, this is another prejudice the developer has, a prejudice that players will always seek the fastest way to play a game. Some players may use WoWhead to quickly become powerful in the game but others may use the system when they are stuck or looking for a specific item in the game. Developers do not always know how players will act. Furthermore, over time World of Warcraft developers have embraced the WoWhead system, even linking to the online system from the World of Warcraft website. The game has been running for so long that any qualms the developer has in regards to players seeing too much of the game's content have become moot. The developer's prejudices have therefore changed as the game has evolved. Interpreting the prejudices of a system thus helps discover why play analytic systems are initially built and help explain why a system evolves over time.

Summary

It is important to understand that play analytics is connected to other forms of game analytics, particularly development analytics. Many game developers are beginning to use, or have used, analytic tools for business, design and debugging purposes. At the same time, game developers are also developing play analytic tools for players (and in some cases players are building their own tools). These two types of analytics ,development analytics and play analytics, are certainly similar; both typically use visualization techniques to present data for example. But, in general, play analytics should be considered different when compared to development analytics. Players are the main audience and play analytic systems need features that enhance a player's gameplay

experience. This means players need to be given access to personal data and systems need to afford different types of data analysis methods, not just the analytic or monitoring data analysis methods. As we move into the next chapter, where I present an in-depth analysis of many more play analytics systems like the ones I presented in this chapter, I begin to describe how play analytic systems are enhancing player's gameplay experience. This includes my analysis of where play analytic systems may be falling short in regards to affording other data analysis methods besides analytical or monitoring methods.

CHAPTER 5

PLAY ANALYTIC CONTENT ANALYSIS

Chapters one through three of this dissertation covers the concepts of data, visualization and analytics. I step through a number of theories, definitions and examples within those concepts, connect them together, and argue how they apply to games. I finally arrive, in part one, at defining the domain of play analytics, a domain consisting of the methods and systems used to interpret data by those who play or spectate a game. In part two, I draw upon the theories I laid out in part one and use those theories to explore current examples of play analytic systems.

In this chapter I present the findings from an in-depth content analysis I performed by reviewing 81 different play analytic systems. Below I discuss the methodology I followed in order to pick and analyze these systems. My final analysis breaks down the domain of play analytics into multiple categories related to common themes or features I found while studying and using the play analytic systems. The characteristics of these categories are discussed one by one and I use the functional and phenomenological theoretical framework laid out in chapter one to interpret how the development and the experience of these systems inter-relate to one another. Finally, I end with a discussion regarding how play analytic systems fit into the ecosystem of digital games in general: what they provide for players, what are their strengths and what are their weaknesses. This analysis leads into chapter six where I discuss the findings from a user study I performed where a number of players used various versions of a play analytic system as they played an game online. The discussion laid out in this chapter has some implications for the results found from the user study, described in chapter six.

Methodology for a Play Analytic Content Analysis

One definition of a content analysis “is any technique for making inferences by objectively and systematically identifying specified characteristics of messages” (Holsti, 1969). The term ‘messages’ in this definition refers to the commonly used information

theory framework laid out by Claude Shannon (1948) where messages represent the content being passed between a sender and receiver along a line of communication. A newspaper journalist is an example of a sender who passes along messages in the form of articles to a receiver, the customers of the newspaper. Content analysis is often used to determine the direction or the attitude of a message or messages between the sender and receiver (Budd et al., 1967; Krippendorff 1980). A typical question answered through content analysis is, “does the sender’s message portray the subject matter in a favorable or unfavorable light?” Analyzing the attitude of a message can help infer what the sender is trying to achieve with their message, for instance it may be helpful in determining how the sender is attempting to persuade the receiver. Other methods of analysis include analyzing the intensity of the message (how favorable is the message to a specific topic?) or frequency of the content within the message (how often does a specific subject appear in the message?) (Krippendorff 1980).

Each form of content analysis must be carried out objectively and systematically; messages must be analyzed in a consistent manner. One way this objectivity is achieved is by building a categorization system and using different categories to organize message content systematically. It has already been established that no categorization system is ever completely objective, nor complete (as discussed in chapter one when covering Bowker and Star’s work on categorization systems (2000)). At the very least, a categorization system can represent a frame of reference, a certain way of organizing content, that attempts to be objective so long as it discloses exactly what it is including and excluding from the content analysis.

As part of my analysis of play analytic systems, I conducted a content analysis using a number of play analytic systems that are currently functioning, or have functioned recently, in relation to a digital game. A broad overview of play analytic systems does not exist and what little work I have done on the subject (Medler, 2009b; Medler, 2011; Medler and Magerko, 2011) only scratches the surface. The content analysis found in this chapter serves as a broad overview of play analytics, compares and discusses common

themes found amongst the 81 play analytic systems I cover, and follows Holsti's procedures for designing a content analysis study (Holsti, 1969):

- 1) Select a sample population
- 2) Create content categories
- 3) Setup relations between categories
- 4) Determine inferences drawn from data

Below I elaborate on each of Holsti's procedures and use them to describe the approach I took for conducting a play analytic content analysis.

Select a sample population

Usually it is customary for a content analysis to analyze a sample population of content from a larger set, due to issues such as limited availability or the sheer volume of content available. If one were to attempt to analyze content from online blogs, for example, typically a great many blogs would have to be discarded simply because the number of available blogs is so large. Play analytics is a much smaller content area, however. While systems such as leaderboards, which are a category of play analytics and are quite common, it is rare to find robust systems that expand beyond simply recording of high scores. Additionally, I have defined play analytics as systems that exist outside of gameplay, making the pool of viable systems smaller. If the external limitation was not in place finding game examples would be much easier. Many games use numerical or "experience" tracking as part of gameplay as part of their game mechanics and rules. Games are defined in this dissertation as 'play with data' and therefore finding examples of how data is used within games is as easy as picking a game at random. Role-playing games, both analog and digital, for instance make extensive use of data by having player characters constantly accumulating data throughout the game's story, often related to the possessions owned by the player character or the character's abilities (Medler 2010). Such game systems parallel play analytics in how data is used and visualized but, as I detailed in chapter one, it is the purpose of this study to focus on systems existing outside of a game space (extra-diegetic systems) in order to understand how players analyze data

when separated from gameplay. The population of play analytic systems studied as part of this content analysis, therefore, exist in some form online, or outside of a game, instead of inside a game while it is being played. There are exceptions to this population rule – systems like Autolog, Ridernet and Battle.net can be used while playing their respective games - but those exceptions are necessary for discussing how play analytic systems are beginning to be integrated inside games.

There are 81 play analytic systems within the content population I study as part of my content analysis. Most systems are attached to a specific game, meaning the game's developer created the system as part of the game's development process and is meant to be used by that game's players. Other systems are built by players who wished to analyze their game data or systems built by third-party companies to aggregate multiple sources of game-related data. I should note that the play analytic systems chosen are mainly systems built around games from the following platforms: PC/MAC, Xbox 360 and Playstation 3. Other platforms such as social games (found on websites like Facebook or Google+) tend to have play analytic features built into the game itself, instead of existing outside of the game as my definition of play analytics demands (which is why they were not included). A similar content analysis conducted by Consolvo using a population of social games found a few related play analytic features, like friend leaderboards or player stat tracking, in many of the social games surveyed (Consolvo, 2011). Likewise, mobile games found on platforms like Android and iOS most likely have the same type of play analytic features as those found in social games (though I have yet to find a content analysis of mobile games). Thus, I chose to focus only on the PC/MAC, Xbox 360 and Playstation 3 platforms because they provided the most ideal play analytic systems to study, those that exist outside of a game and function as independent systems. Additionally, given that the systems covered in this study are separated from the game's they are built for, nothing is stopping a developer from creating similar systems for games on the other platforms not covered. A developer of a social game on Facebook could build a similar play analytic system as the ones found within my study's population of systems. While I do not necessarily argue that all of the play analytic categories discussed in this chapter can apply for any play analytic system built for any game, I do

argue that many of the categories have a possibility of being relevant to a number of games on a number of platforms. For a full list of the analyzed play analytic systems, see Appendix A. [REF]

The functional properties that define categorization systems, and thus play analytic systems, act as collections of categories in this analysis. Since the features or functions available in a system influences how players use the system, the content analysis begins with categories related to the comparability, visibility and control properties. The phenomenological properties (interconnection, exposure and historic) are harder to categories as part of a content analysis because this analysis only takes account of a system itself, not the player using a system. Instead, the forth procedure of this content analysis uses the categories created around the functional properties to help infer the player's experience and explains how the phenomenological properties are related to the functional content categories.

Create content categories

An initial investigation of the play analytic system population, before the final content analysis, was conducted in order to determine the major categories for defining the breadth of common features found within play analytic systems surveyed. This investigation consisted of myself logging into or otherwise using each system, one by one, to determine every feature available. Features are defined as capabilities that compare data, make data visible or control the administration of data, which relate to the three functional properties found in the functional/phenomenological framework. As discussed when defining the three functional properties in chapter one, the functional properties are meant to define how categorization systems are built and maintained. The properties can therefore be used to define any features that exist within a play analytic system. For instance, a friend leaderboard (a leaderboard that only consists of scores related to a player's friends) is a feature because it compares players, makes certain game variables visible, controls which variables are visible and controls which players are allowed on the leaderboard. But a friend leaderboard is just one type of leaderboard, a leaderboard with a specific feature set. The initial investigation I conducted provided the

means of finding common, and uncommon, features of a play analytic system as well as the overarching categories, or groups, that those features fell under. Leaderboards, for example, is a play analytic category, while various features of leaderboards make up the variations within that category.

Eight play analytic categories were decided upon based on the initial investigation of the play analytic systems. Categories can be thought of as design patterns. Within each category are variations that represent features found at least once, often more than ten, system and consist of common attributes to define them as a distinct feature within the overall play analytic category. A friend leaderboard, for example, is a variation of the leaderboard category. Each of these categories are placed into two overarching themes: gameplay and creation. These themes make it easier to discuss how the categories relate to each other based on what they provide within a play analytic system. The eight categories, organized under their themes, are defined below but a detailed analysis of each category, and the variation found within each group, is also provided.

Gameplay

Categories under the gameplay theme represent play analytic patterns that specifically use data collected from a player's gameplay. These categories may include other data, facts from the game's environment perhaps, but generally focus on providing access to a player's gameplay data. The following gameplay categories are:

Leaderboards – A list of players ranked in order based on a set of criteria. A score earned during a game is an example of the type of ranking criteria; players earning higher scores are ranked higher than other players.

Player Statistics – List of variables, calculations or other data related to the gameplay of a single player. It should be noted that the term 'statistics', which is often abbreviated as 'stats', should not be confused with the actual science of statistics. Some play analytic systems use no statistical methods at all beyond the accumulation of data points and reporting frequency of gameplay event occurrences. The term "signs", as in

life signs, is another appropriate word to use in place of stats but “stats” is the normative term.

Groups Statistics – List of variables, calculations or other data related to the gameplay of a group of players, often associated with a group identity chosen by the group itself.

Match Statistics – List of variables, calculations or other data related to a single instance of a played game, known as a match or contest. Often used in games where players most likely play many “matches”, a self-contained contest, over the course of their time playing the game. A chess match is an example of how matches are perceived, one where two players oppose each other attempting to win the game but after concluding the first match a new, second match can begin.

Global Statistics – List of variables, calculations or other data related to the gameplay of an entire player population.

Maps – A visual depiction of a game’s environment and geography that may include interactive features and display data related to the game or a player’s gameplay.

Content Creation

Categories in the creation theme represent features which deal with data created for use in a game but does not necessarily include data recorded from a player’s gameplay. Data can be created while a player is playing but the player is purposefully creating the data being recorded. The following creation categories are:

Content Generation – Features that allow players to create content (characters, artwork, levels) to use in a game or features for sharing content that are created in a game. Generated content is generally considered to be infinite because players can continuously produce more content, so long as there is the means to store the content and players continue to create the content.

Content Database – A type of database containing data related to the content created as part of a game’s development process (quests, items, enemies) and excludes data created as part of a “generation” process. Content data is often created by the developer but players can create additional content data related to their experience playing the game (for example, create a walkthrough detailing the best strategies for finishing a level).

Setup relations between categories

Each category within a theme represents a different exclusive feature but most of the play analytic systems analyzed have features from more than one of these categories. This means the easiest relation between the categories is merely the relation of availability. Play analytics systems have many features and are thus marked as containing many of the nine categories. Systems with ‘leaderboards’ and ‘player stats’ are marked as having both features. Further relations between categories exist when features begin to combine within specific systems. Maps can appear alongside player stats or leaderboards in group stats. Detailing the relations between the categories, along with their variations, is part of the analysis provided in each individual categories analysis below. [REF Appendix A] provides a table listing which groups exist within each system.

Determine inferences drawn from categories

As the final content analysis procedure, determining inferences from the content categories includes making the connections between the functional categories and the phenomenological properties of play analytic systems. As I step through each play analytic category below, describing the variations and giving examples of each, I also provide arguments for how each category can be interpreted by the six properties found in my functional/phenomenological framework: visibility, comparability, control, exposure, interconnection and history. Therefore the analysis of each play analytic category is laid out as such:

- 1) ***Overview*** – An introduction and definition of the category.

- 2) ***Variation List*** – The different feature variations are listed for each category.
- 3) ***Category Examples*** – Systems that use the category variations are described.
- 4) ***Theoretical Analysis*** – The category is interpreted based on the six functional and phenomenological properties.
- 5) ***Summary*** – Concluding remarks and critiques about the category.

These long descriptions of each play analytic category are meant to describe the common features of play analytic systems in order to highlight design issues and discuss how designers, and players, can mold play analytic systems to match their vision of the system's purpose, or in the players case mold the systems so provide something valuable to their gameplay.

Below I begin by covering the play analytic categories found within the gameplay and creation themes. Generally, the way I step through the categories is based on how common or dependent each category is based on the other categories. For example, leaderboards are first because they are highly independent compared to the other categories. Leaderboards often existing entirely on their own and devoid of other play analytic categories. A category like match statistics, however, regularly depends on player and group statistics because match statistics tend to relate to the stats collected for individual players or a group of players. Although, the choice of how to order the categories is my own way of organizing them and the dependencies implied are not definitive relations. A play analytic system offering match statistics does not necessarily have to offer group statistics for instance. After covering each play analytic category I finish with a final discussion where I describe some common threads through a number of the play analytic categories.

Leaderboards

Leaderboards are perhaps the most ubiquitous form of the play analytic categories. Since the mid-1980's they, along with high-score concept, have appeared in a countless number of games (Medler, 2009b). This is perhaps due to the ease of which a leaderboard can be setup for a game, especially today when platforms like Xbox Live

provides stock leaderboard systems developers can use. The steps to create a leaderboard are as such: (1) choose a gameplay value, or some set of criteria, that can be used to order players (perhaps a player's final score during a game), (2) have each player submit the value at a certain point during their gameplay experience (often at the end of the game), (3) the player with the better value (whether it is higher or lower than other scores) is placed above any players with worse scores. Basically, this just means a leaderboard defines which player had a better score. Players with better scores are ranked higher than other players and are acknowledged as being better players (at least according to the leaderboard criteria). The act of ranking players, and recording their values, on a leaderboard instantaneously promotes competition in games, especially when players cannot play concurrently or right after each other. Leaderboards allow players to compete across time. This is one reason why leaderboards are so ubiquitous across digital games.

Half of the play analytic systems studied 43 out of 81 have some sort of leaderboard variation. This number is not as large as the number of systems with player statistics features (which I cover next) but my choice of systems may be the cause of this. If I were to include an analysis of every game with a leaderboard, I would have an impossible population to study; there are simply too many games with a leaderboard. Most leaderboards are only accessible while playing a game, usually as a menu option, and are not available outside of the game. Dead Space 2 for example has a few of the leaderboard variations I describe next but Dead Space 2 does not have a play analytic system external to the game.

The leaderboard variations discussed next mainly cover how players are ranked and what type of players are ranked. Some leaderboards only rank players based on a single value while other leaderboards may rank players uses five or more values. Other leaderboards only rank a limited number of players, hoping to make the leaderboard more prestigious for those players who are fortunate enough to be ranked. Finally, some leaderboards may rank only a player with their friends or may rank players based on the groups they play the game with. However, each of these variations follow the same

model for all leaderboards: set a criteria, gather the criteria from players and order the players accordingly.

Leaderboard Variations

The following design variations of player statistics appeared in the play analytic systems surveyed:

Single Criterion Leaderboard



STEAM LEADERBOARDS		
Leaderboard: The Hallway		
THE HALLWAY		
#1	dArKneSS x DrAgoN	42,029,854
#2	Jason Switcher	37,516,903
#3	HODJA	33,019,502
#4	Closclu	28,510,248
#5	Avaluque	15,020,567
#6	Azrael	3,033,007

Figure 5.1: The single criteria leaderboard for the game Orcs Must Die. A players final score in the game’s different levels is the single value that rank each player on the leaderboard.

Players are ranked according to a single piece of criteria, be it a score, variable or other ordinal value. A single criteria leaderboard is the “ideal” example of a leaderboard. The very first leaderboards created were designed to rank players based on the highest scores achieved in a game. Today, single criteria leaderboards are common but there is rarely only one leaderboard per game. Most games now have a leaderboard for every level or game mode, even if they are still rely on a single criteria model. The leaderboards for Orcs Must Die, for example, are single criteria but there is a leaderboard for each

level in the game (REF see Figure 5.1). A player in Orcs Must Die can be ranked number one on one level's leaderboard but fortieth on another.

Multiple Criteria Leaderboard

Players are ranked according to multiple criteria values, be they scores, variables or other ordinal values. Instead of ranking players based on a single criterion, multiple criteria leaderboards rank players along multiple values on the same leaderboard. These leaderboards often allow those viewing the leaderboard to reorder the rankings based on one of the multiple criteria present. Leaderboards for the game MAG allow a viewer to order the leaderboard based on: experience points, kills, deaths, kill/death ratio, headshot kills, most valuable player award, wins, losses and win/loss ratio (REF see Figure 5.2, the system has since been decommissioned). All of the values were displayed at the same time on the MAG leaderboard but only one value is used to rank the players at a time. Also, in the same way there may be many single criterion leaderboards for the same game, most play analytic systems with multiple criteria leaderboards have more than one. Common separations between multiple criteria leaderboards are geographic region, time period and platform (for games that appear on more than one gaming platform).

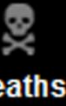


Rank	PMC	Player Name	XP 	Kills 	Deaths 
1		BiGBOSS_UGUR	15752184	317936	153864
2		Jamesamundo	14758887	186304	206623
3		zuena13	12049813	278095	105362
4		vadik1990	12018820	212264	154996
5		CERVEZA2001	11593572	238787	93702

Figure 5.2: The multiple criteria leaderboard for MAG ranks players along many dimensions including the players total experience points earned, kills made and deaths.

Limited Leaderboard

Only a certain number of players are allowed on the leaderboard at any one time. In a similar way single criterion leaderboards are seen as the “ideal” example of a leaderboard, a limited leaderboard also fits that description. Early leaderboards on arcade machines had limited ranked slots, usually for ten players, because the machines had little data storage capacity. Many current online leaderboards do not have this limitation and can place as many players on a leaderboard as they want. However, there are still leaderboards that limit their player numbers. Super Crate Box for example has a few different leaderboards, for different game modes, but only allow ten players on each board. As part of the Starcraft 2 leaderboards on Battle.net only 200 players are allowed on the grand master leaderboard, symbolizing the top 200 Starcraft 2 players in the world (REF see Figure 5.3). Limiting the number of players on a leaderboard can make the boards more prestigious because only the best players are represented.




	Rank	Player	Points	Wins	Losses	Race
	1st	 GoSuostojy	690	65	15	Zerg
	2nd	 oGsVINES	680	75	16	Protoss
	3rd	 MVPTAILS	618	88	33	Protoss
	4th	 Fenrir	598	210	30	Zerg
	5th	 More	535	100	35	Protoss

Figure 5.3: Starcraft 2 limits the number of players on each leaderboard. For example, the Grandmaster leaderboard (pictured) only allows 200 players to be ranked.

Friend Leaderboard

Only friends of a player are displayed and ranked on the leaderboard. Friend leaderboards are a way of having a limited leaderboard that everyone significant to a player can appear on. Instead of limiting the players based on the global population of

players, a friend leaderboard limits the players to a group of friends. Many play analytic systems allow players to mark which players are their friends. Using that data, a friend leaderboard only displays the scores or values for a player's friends, creating a limited leaderboard that is uniquely relevant to a player. The Autolog and RiderNet systems built into Need for Speed: Hot Pursuit and SSX, respectively, are built around the concept of friend leaderboards. In these systems, each player's score is ranked against their friends. If the players achieve better score than their friends the player can earn more experience points, thus enticing players to continuously try to beat their friend's score. Autolog and RiderNet can also provide recommendation listing which races a player should play based on how close a player's score is to the next highest friend's score (REF see Figure 5.4). A closer score means a player may have a better chance of beating their friend. However, friend leaderboards do not necessarily work if a player has no friends playing the specific game the leaderboard is built for. Without any friends in the game, a player is constantly at the top of a leaderboard with no one challenging them. Some play analytic systems get around this by offering friend leaderboards in addition to other, global leaderboards that include the entire player population.

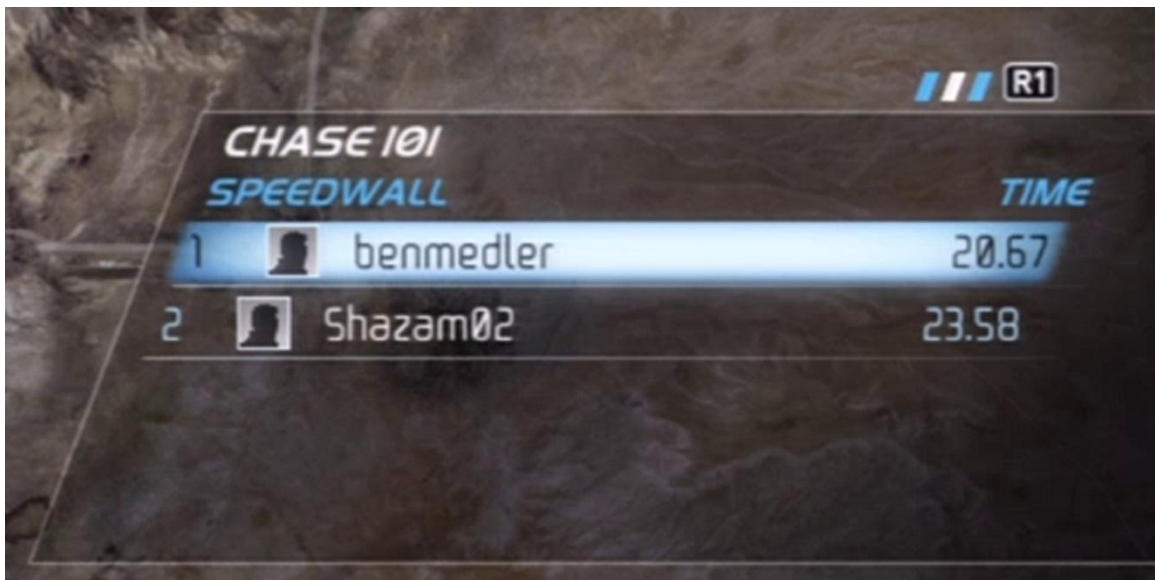


Figure 5.4: Autolog has a friend leaderboard for each race in the game Need for Speed: Hot Pursuit.

Only players designated as friends can appear on a player's friend leaderboard.

Group Leaderboard

A leaderboard ranking groups of players instead of individual players. When players can form groups and participate in a game as representatives of that group, whereby they receive scores based on their group's performance, a system often ranks these groups using a group leaderboard. Starcraft 2's leaderboards on Battle.net and the World of Warcraft arena battle leaderboards on the WoW Armory are both instances of group leaderboards (REF see Figure 5.5). Any time a set of players, usually ranging between two to five in the case of Starcraft 2 and WoW, form a group and fight in a match they are ranked on a leaderboard as a single group. Players may form many groups at the same time in these systems as well. A Starcraft 2 player may have many two, three or four player teams they are attached to, with each team being ranked on their respective leaderboards (based on team size). Other games such as Killzone have groups fight in tournaments, where each tournament or "Season" (as they are called in other games that offer tournaments like FIFA 2012) has its own respective leaderboard.



Top 2000 Arena 5v5 - Bloodlust

Team: Realm: Faction: Rating: -

☒ Show all team compositions ☐ Filter by team compositions

Showing 1-50 of 1986 results

#	Team	Realm	Faction	Wins	Losses	Rating
1	Jido cant keep up	Barthilas		143	2	2668
2	VIP Pets US	Tichondrius		304	19	2631
3	Tactful Damage	Tichondrius		87	6	2481

Figure 5.5: World of Warcraft Arena leaderboards rank groups of players, each group consisting of between two and five players, based on their performance in player versus player combat.

Leaderboard Examples

One example of the quintessential, or “ideal”, leaderboard is the leaderboard system built for the game *Orcs Must Die*, a real-time tower defense game. The game has multiple levels that a player progresses through over time and all the levels are 3D environment (other tower defense games are played on a 2D plain). The goal of the player is to stop a series of enemies, the orcs, from reaching one side of the level (the orcs often flood into the level from the opposite side of the level). Players can lay traps and ‘towers’ in order to help them attack the enemies as they run by. Each enemy killed gives the player more points and killing enemies in certain ways, e.g. killing many enemies at the same time, can earn players even more points. After completing a level, every player is ranked on a leaderboard based on the number of points they earn while playing the level. The *Orcs Must Die* leaderboard available on the game’s website, which includes other player statistics, is a type of single criterion leaderboard and a limited leaderboard. Players are only ranked based on their final point scores and only fifty players are shown for each level’s leaderboard. Conversely, while playing the game players can see their rank among every score ever submitted, not just the top fifty players. The in-game leaderboards also have a friend leaderboard option that does not appear on the *Orcs Must Die* website. In this case, the play analytic system (the one available on the *Orcs Must Die* website) has less leaderboard variations than the leaderboard offered in the game.

A different leaderboard system compared to the *Orcs Must Die* leaderboard is the system built for the game *Need for Speed: Hot Pursuit*, a competitive car racing game. Named Autolog, this leaderboard system is integrated into *Need for Speed: Hot Pursuit*, making it an edge case concerning the definition of a play analytic system, and system exclusively uses friend leaderboards. *Need for Speed: Hot Pursuit* has a number of individual races players can take part in and each race has a friend leaderboard attached to the specific race. Autolog is setup to monitor each of these individual race leaderboards and provide recommendations to players telling them which races they should try next. For example, if a player’s score was beaten by one of their friends, Autolog may recommend the player play the race again to try and beat their friend’s score. Using friend leaderboards makes it easy for players to keep track of how well they are playing, because there are fewer players to track, but the Autolog recommendation

system also makes it easier for players to find races where their scores are close to their friend's scores. Unfortunately, if a player has no friends playing Need for Speed: Hot Pursuit the Autolog system becomes impotent and is not able to provide recommendations. While other systems offer multiple types of leaderboards, like the in-game Orcs Must Die leaderboards, Autolog's major flaw is focusing solely on friend leaderboards.

Functional and Phenomenological Aspects of Leaderboards

Below I discuss how leaderboards can be interpreted based on the six properties found in my functional/phenomenological framework. While the functional aspects of leaderboards may not seem very complicated, how leaderboards have progressed over the years has made them much better at connecting players together.

Visible

The data made visible through leaderboards typically only cover high level details about a player. Data such as high scores or most kills are used to represent the player. Most of the other data describing how the player achieved a high score or killed the most enemies is lost and made invisible by leaderboards. A player's rank amongst a group of players is made visible too, not just their score, but the group of players the rank is based on can change based on which leaderboard variation is being used. On a friend leaderboard, the scores and ranks are only made visible to a player's friends. On limited leaderboard only the players with the best scores are made visible. Compared to the other types of play analytic features leaderboard does not make much player data visible and seems to make only the best players visible to the community at large.

Comparability

The act of comparing is fairly regulated within a leaderboard, sometimes based on a single score or value, and is used to promote competition. Each board has a set criteria by which each player is compared and ranked against. Whether a single player, a group of players or a player's friends are being ranked, they are always ranked in the same way given the criteria of the leaderboard. Thus, one can argue that comparing players is the

major service, if not the only service, a leaderboard provides. This is why when a leaderboard only has one player it cannot really be defined as a leaderboard because it has no players to compare. Friend leaderboards, therefore, run into a problem when players do not have any friends in a game, the leaderboard cannot rank a single player against themselves. More than one player is needed to be compared on a leaderboard. Otherwise, the leaderboard cannot function as a system to help promote competition between players.

Control

Leaderboards tend to be officially maintained by a game's developer. Developers set the criteria for how players are ranked on a leaderboard and keep records of the highest scores. They also hold the power to reset a leaderboard, clear out all of the scores, or remove players they feel should not be on the leaderboard. The leaderboards for Halo Wars were reset a few months before the game's play analytic system was threatened to be shut down. A reset can be beneficial to players who continue to play the game because older scores from players, who may not be playing the game anymore, are removed and current players are given the chance to attain a higher rank. Sometimes players also find ways of hacking a leaderboard, to record scores that are impossible to achieve in the game. When an instance of hacking occurs developers may take steps to remove the offending players scores from the leaderboards. In a four month period while watching the Orcs Must Die leaderboards, as part of this research, many top scores were removed from the leaderboards most likely due to players find ways of cheating or hacking the game. The decisions to alter a leaderboard tend to fall with the game's developer, or whoever runs the leaderboard.

Exposure

When viewing a leaderboard two sets of data are exposed to players. This first is the player's name. Every leaderboard displays the players name or gamertag. If leaderboards are meant to promote competition then players need to know who they are competing against. The second set of data exposed to players is the criteria the leaderboard ranks players with: scores, kills, deaths, play time, etc. This second set of

data exposed gives players a sense of acknowledgement, that they have achieved something within the game. Too often, however, this sense of acknowledgement goes unnoticed by other players.

If a player is not in the top ten, let alone the top 200 like on Battle.net's Starcraft 2 grand master leaderboard, then other players probably couldn't care less about their score. Even though a player's name and scores are being exposed, leaderboards are often too flooded with players for anyone to stand out. Leaderboards found in COD: Elite can have over ten million players, for instance. What player is going to take the time to look over everyone's score? Players that sit at the top of heavily populated leaderboards may find meaning from the leaderboard, they may see it as a validation of their gameplay abilities, but for everyone else the leaderboard becomes less meaningful. Unless a player is ranked at the very top of a leaderboard their data is less exposed because leaderboards make worse scores more inane to other players. There are just too many scores to keep track of. That's why friend leaderboards are argued to offer a better way of promote competition; fewer players on the board means players begin to care about everyone's exposed score and try harder to beat the next highest score. Exposing scores on leaderboards with fewer players makes the scores more meaningful within the context of smaller groups.

Interconnected

A leaderboard on an arcade machine used to have the power to connect players across time while staying in the same space (Medler 2009b). Any player with a ranking on an arcade machine's leaderboard knew that the other ranked players on the board also played at the machine. The ability for a leaderboard to remain in one space doesn't really exist anymore, outside present day arcade machines. Online leaderboards cover the whole world. Players, therefore, cannot connect to each other based on space and must find other ways to connect to one another using leaderboards.

Friend leaderboards seem to provide the best example of how leaderboard can still connect players together even though players are separated by time and space. They

allow players who know one another to post their scores and only their scores are compared against one another. When compared to other leaderboards that list every player's score, all over the world, the intimacy of being compared against a friend's score makes friend leaderboards a better way for players to connect and compete against each other. But there is something to be said about the power of being on top of a leaderboard, above the masses of other players. Anecdotally, as I analyzed the systems for this study, I found that many players on top of a system's leaderboard tended to have a large number of followers or fans (at least on the systems that had such a feature). Being recognized as one of the top players is another way for players to connect with one another. Even if the other players only become fans of those high performing players, being a spectator of a good player is really no different from a spectator going to see a baseball or football game. Leaderboards allow players to connect with the best players, whether they are ranked amongst a large group of players or a close circle of friends.

History

Leaderboards exist in the moment. Rarely have I seen a leaderboard that shows a player's score progression over time or even list when a score was ranked (and I saw neither instances in the systems I surveyed). The player at the top is the player at the top, it doesn't matter if they played the game last week or last year. That's why acts such as leaderboard resets or the removal of players can go unnoticed. There is no history of where the scores came from and what players have been on the board across time. Sometimes there are leaderboards that have a "previous ranking" field that lists what rank number a player was yesterday or last week but that is as much historicity as most leaderboards can muster. While leaderboards work fine as a means of promoting competition, especially friend leaderboards, there is room for some exploration into how leaderboards could be made more historic. Having a more historic-centric approach may make leaderboards more relevant, and enjoyable to monitor, especially for e-sport related games where there are many spectators following the top players in the world.

Leaderboard Summary

The main function of any leaderboard is to compare and rank players. As a result, players are spurred into competition. This can benefit games with competitive aspects and as the leaderboard variations show there are many ways of approaching player comparisons using a leaderboard. Sometimes single scores or criteria are used to rank players while other times players are ranked along a set of multiple criteria. Having more than one criteria can help cover different aspects of gameplay players find more meaningful than other aspects, for example one player could value their accuracy value higher than their win/loss ratio. There may also be separate leaderboards for different levels or game modes, effectively spreading the leaderboards out so competition is compartmentalized to specific areas and more than a few players can achieve high scores. Other leaderboards limit the type of players compared, sometimes only showing the top players in the world or limiting the leaderboard to a player's friend or group.

Using different leaderboard variations also alters how leaderboards are perceived by players. On a global leaderboard where every player is ranked along a single, or multiple, criteria, players become lost in the endless list of scores. A player may be ranked '200,000 out of 700,000' players. That is not a ranking, it is another inan number in the game, virtually meaningless to both the player and every other player. The value is too high to make a player feel good about their performance. If a player continues to play they may increase their rank substantially but unless they make it to the top 100 players there rank remains inan. In order to get around this fact, variations like friend leaderboards are used to make it easier for players to compete solely amongst their friends. Even if players have hundreds of friends, which means a player could have a very low ranking, being ranked with friends enriches the leaderboard values. Players care more about competing with their friends than a global population of strangers (unless the player is ranked high enough on the global leaderboard to be recognized as one of the best players in the world). Being partitioned into a group of friends on the leaderboard also affords the ability for players to talk with their friends that have higher rankings, perhaps learn how to increase their ranking in the game. However, the drawback of a friend leaderboard is revealed when a player has no friends playing the particular game using the leaderboard. The player is left alone with no one to compete with. Their ranking

becomes inan again, they have no baseline for judging whether their scores are high and must only compete with themselves (which tracking one's performance over time can be fulfilling but should be considered player stats). Last, there is rarely any historic perspective when viewing a leaderboard. Players are often given no information regarding when a score was set and how players have progressed over time. Leaderboards are sometimes separated based on weekly, monthly or 'all time' scores, therefore separating the score of players who are currently playing from the older players score, but the perception of leaderboards is typically ahistoric and leaderboards are often seen as static.

As we continue on to review other play analytic categories, leaderboards are often associated or mixed into those categories. Leaderboards often appear in match stats, alongside global stats and can also be used to rank player generated content. They are the most prolific category of play analytics and have a great influence on how players are ranked and compared in other categories.

Player Statistics

A common feature of most online systems is the user profile, a representation of a user within the context of the system. Online social networks like Facebook or Google+ allow users to create profiles and those profiles act as hubs for each user's activity on the social network. Other online systems follow the same model; communication services (Skype, AIM), wikis, blogs, forums, content sharing websites (YouTube, Vimeo), e-commerce services (Amazon, Ebay) and any number of other online systems provide users with the ability to make a user profile. Online systems related to games, e.g. play analytics, are no different. Every major game publisher (2K, Activision, EA, Ubisoft) has an account system where users sign up, get profiles and are given access to various game websites or online content. Gaming platforms (PSN, Steam, Xbox, not to mention other mobile platforms like Android and iOS) also have a similar setup where players create gamertags and are given user profiles. The act of giving users the ability to construct an online persona by using a user profile is prolific. There is one key difference, however, when discussing user profile systems in respect to the "profiles" given to players within a

play analytic system. User profiles allow users to create and add their own content to their profile while other “profile” systems, especially the ones found in play analytic systems, automatically add content to a user’s profile with very little user intervention. Automatic profiles, like those found in play analytic system, act more like dossiers or compiled reports than actual user profiles. And the content found within those dossiers is often referred to as player statistics.

More than two-thirds of the play analytic systems covered in this content analysis have some form of player statistics. The reason so many play analytic systems have player statistics is because play analytic systems are built for players to use. Players play a game, they generate data and the data is turned into statistics. The statistics are displayed through a play analytic system and, hopefully, the system keeps the players engaged enough to keep them playing the game. There are of course other data sources, e.g. game assets or maps, but, generally, players are both the users of play analytic systems and the main data creation source for the systems. Player statistics are the numbers, attributes or other data generated by players while playing a game. A player’s fastest time in a race, the number of enemies they have killed, the achievements they have earned, etc. are all collected and presented as player statistics. Anything related to an individual player’s gameplay, which generally means most data a player generates while playing a game, is considered player statistics.

Player statistics present a player’s data as part of a dossier report and acts as a representation of how that player has acted while playing a game. For example, player statistic variations like activity feeds include a constant stream of gameplay events happening during a player’s gameplay, activities such as: winning achievements, joining a player group or finding a game item. Other player statistics examples found on a dossier include presenting values describing how well a player is performing, like high scores or win/loss ratio. What this means is players have little control over what statistics are collected and shown on their dossier pages, which is why I refrain from using the term player or user, profiles. The player statistic variations presented below reflect this lack of control. Many of the variations cover features related to collected data from users

passively with only a few variations (e.g. friends list and media sharing) allowing users to actively add data to their dossiers.

Player Statistic Variations

The following design variations of player statistics appeared in the play analytic systems surveyed:

Accumulated Statistics

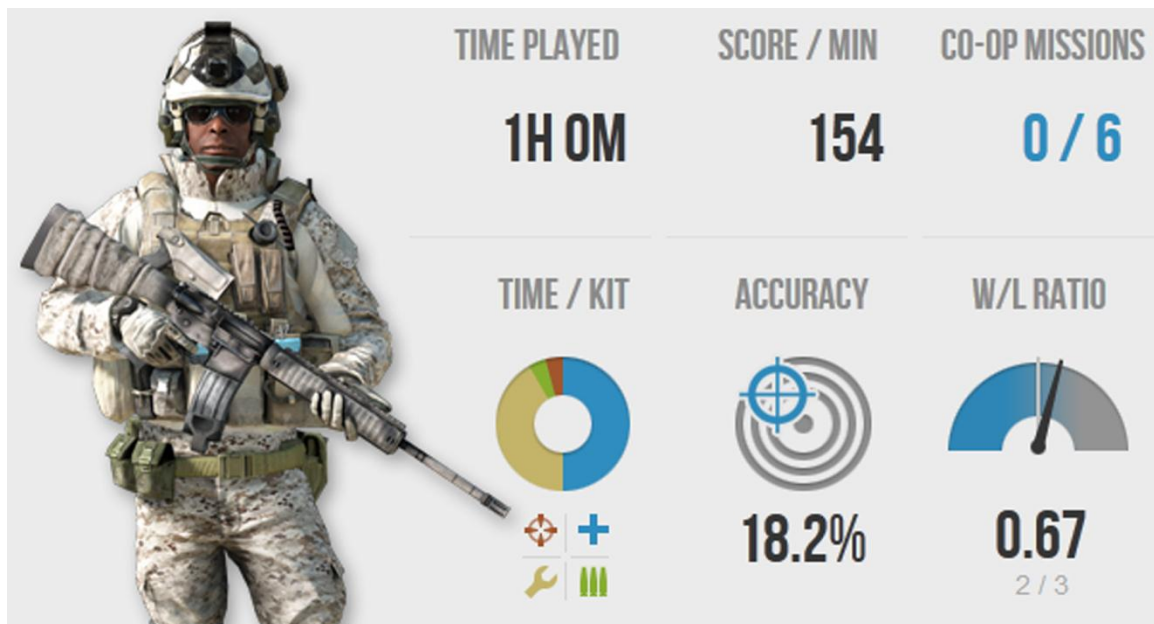


Figure 5.6: Battlelog displays a number of accumulated statistics on a player's dossier.

Values counting gameplay event occurrences or other gameplay related values are accumulated into a set of statistics for a single player. Almost all, if not all, play analytic systems with player statistics have accumulated stats. First person shooter games provide the total number of kills a player has performed and how many deaths a player has experienced, for example (REF see Figure 5.6). Other accumulated stats may include: total time spent playing a game, how many times a piece of equipment is used, how many special moves a player performs, the total number of earned experience points, etc. Also accumulated stats may accumulate forever. Having bloated player stats, where a player has accrued huge values over time, is a common occurrence for older play analytic systems and can be a problem if the system is not designed to visually represent larger

values (graphs or text may visually appear wrong in the system if a value grows too large, for instance).

Progress Statistics

Similar to accumulated stats, progress statistics are values which accumulate until a threshold or final value is reached. Accumulated statistics may accumulate forever but progress stats eventually stop. A basic example of a progress stat is the overall game completion percentage, which means a player completes every aspect of a game (usually resulting in a 100% completion). As player's progress and complete activities or levels in the game their completion progress percentage goes up (REF see Figure 5.7). Finally completing everything in a game means a player's completion progress stops at 100% even though they may continue playing the game. Other progress stats include: unlocking abilities, player experience level (which usually has an upper threshold) and achievements with multiple requirements. As with accumulated stats, progress stats are very common among play analytic systems that offer player statistics.

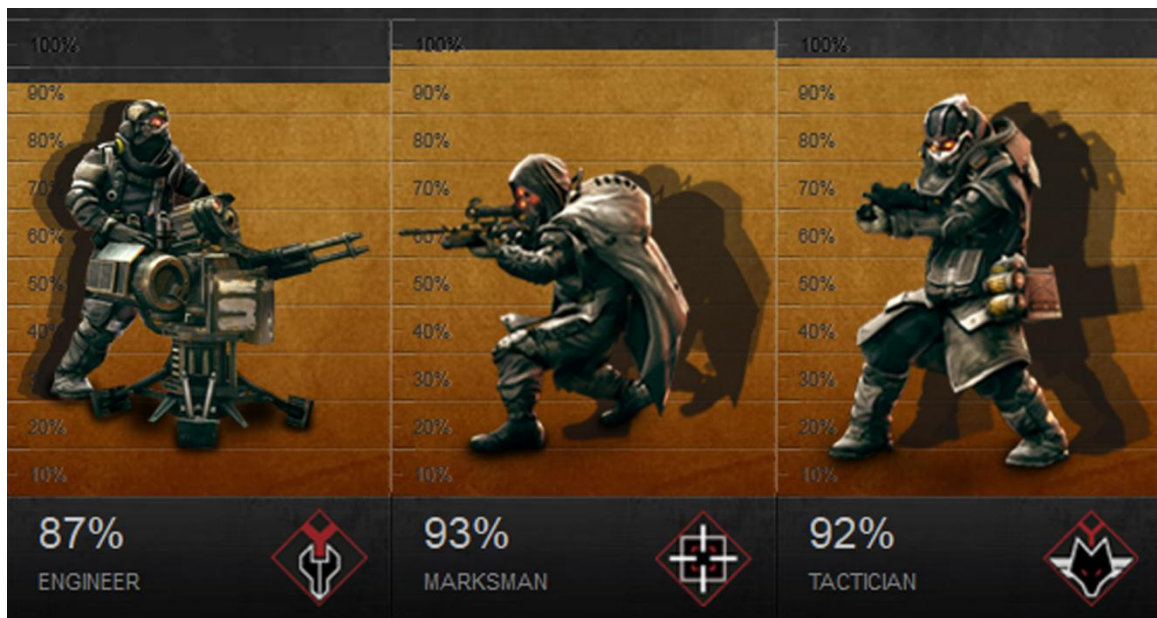


Figure 5.7: Progression along each of Killzone 3's player classes is displayed in each player's dossier.

Game Mode Statistics

Games with multiple game modes may separate player statistics based on the modes. Single player and multiplayer modes are the most common separation. Single

player modes typically consist of a campaign or other story driven experience, that the player plays by themselves. These experiences are usually different compared to their multiplayer counterparts. Multiplayer modes are routinely matched based and each multiplayer match can last for a shorter amount of time than a game's single player mode (REF see Figure 5.8). Therefore, the stats collected from both modes often reflect these differences. Stats from single player modes may focus on the story aspects of the game, looking at how a player completes the objectives set by the campaign, whereas multiplayer mode stats may focus on how players interact with each other, e.g. kill/death ratios or awards for actions during individual multiplayer matches. Besides single player and multiplayer modes there are also: co-op mode (multiple players playing as a single unit, often playing in a single player mode), modes created from game expansions (Spore has a regular game mode and an adventures mode that was added through an expansion), and modes that are entirely different games (EA sports combines many sports titles into on centralized player hub). Game mode statistics often appear in play analytic systems when there are distinct modes and these modes appeal to different player groups (for example, some Call of Duty, Battlefield or Starcraft players only play the multiplayer mode and disregard the single player mode).



Figure 5.8: Player stats can be broken up based on different game modes. Starcraft 2 has a number of match modes such as Free For All, Custom Games and Ranked modes.

Demographic Data

Data related to a player's actual identifiable demographic data. A player's real name, gender, geographic location are all examples of demographic data some play analytic systems provide as part of player statistics (REF see Figure 5.9). Other data such as email addresses and usernames from other communication services (i.e. Google Talk or Instant Messenger) also fall into demographic data. Play analytic systems do not regularly give out demographic data and almost always give the user the ability to release as much of this data as necessary. The ability to control the release of demographic data is mainly due to privacy concerns but, also, demographic data usually has no relation to the overall player stats being collected by a play analytic system (a player's gender doesn't affect their win/loss ratio), making the data meaningless to other players (except for cases where geographic location may link to real world factors such as support for local sports teams, as found in EA sports player hub system).



Figure 5.9: EA Sports Active provides limited demographic data within each player's dossier, including: location, age, gender, height and weight.

Load-out

In games where players have a consistent set of equipment or abilities they use, a load-out describes the equipment or abilities a player is using. Games that give players the ability to select and modify different aspects of their avatar typically provide a load-out. This variation provides players with the ability to show their diversity in a game

based on the equipment or abilities they have chosen. First-person shooters and role playing games often have load-outs for describing the type of weapons and armor a player is using. Each weapon and armor piece, or other item, has a set of stats or abilities related to gameplay, and those stats are presented as part of the load-out (REF see Figure 5.10). Other examples of load-outs include: skill trees, avatar clothes or modifiable vehicles (usually found in racing games).

EQUIPMENT			
Main		39	Demise Tome
		39	Demise Jewel
Off			
			
Armor		1	Perento's Tunic
		1	Perento's Pauldrons
		42	Researcher's Leggings
		38	Justice Gloves
		33	Expert Sobi Shoes

Figure 5.10: Load-outs in Aion's player dossier provide a list of the items a player's character currently has equipped in the game.

Reputation and Ranking

Players are given a reputation or rank based on their behavior. A player's level, based on experience points, is a common ranking motif but some play analytic systems rank players based on how they perform in the system itself. Newgrounds, for instance, ranks users based on how often they rate games, write reviews and write forum posts (REF see Figure 5.11). These activities contribute to a player's reputation and ranking within the system, giving users with a better ranking more weight when it comes to rating games or other content available on Newgrounds. Other reputation and ranking values for an individual player may be related to creating content, however ranking player generated content is discussed later in the Content Generation section in this chapter.



Figure 5.11: On New Grounds, a web portal for sharing Flash-based content, users are given ranks based on how they rate and vote for games that appear on the website.

Compare Players



Figure 5.12: The Assassin's Creed Web Battle system allowed players to compare their player stats and players earned points for having higher stats.

The ability to compare the player statistics of two or more players. Comparing players is often a feature of play analytic systems made for competitive games. Each player's accumulated and progressive stats are compared against each other to see which player has better values (sometimes this means a higher number, other times lower). Assassin's creed 2's play analytic system had a 'web battle' comparability feature where

players earned points for having better stats than the other player they were ‘battling’ (REF see Figure 5.12, the feature has since been removed). Sometimes a system may encourage comparing stats between friends and other system encourages comparing stats to other players (compare buttons may be shown by each player’s name on a leaderboard, promoting players to compare themselves to the top players).

Trophy Rooms

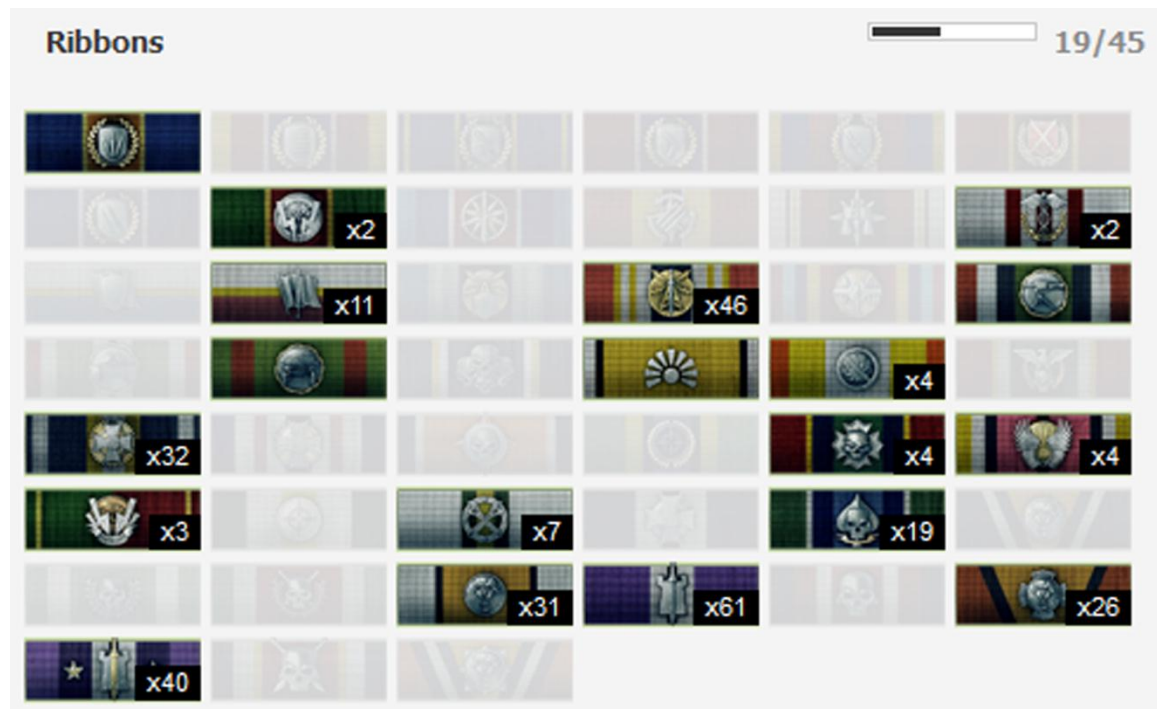


Figure 5.13: Many player stats include a trophy room area to show off the awards, medals or other trophies a player has earned (pictured is one example of a trophy room from the Battlelog system).

A list of any achievements, awards or trophies won by a player. Achievements, awards, trophies, etc., represent single items that acknowledge something related to a player’s behavior. Some of these earned items may relate to accumulated stats but usually work more like progress stats. A simple achievement can be earned by starting a game for the first time or a trophy may be earned when a player is ranked number one on a leaderboard. The term ‘achievement’ often refers to one-time items, meaning after a player is awarded an achievement they have the achievement forever but may never gain the same achievement again. Other monikers are used for items that can be awarded multiple times: trophies, awards, ribbons, medals, etc. For example, killing ten enemies in

a multiplayer match may be an award given to a player every time they accomplish that feat in a match. The term ‘trophy room’ describes how these earned items are often put on display as part of a player statistic dossier (REF see Figure 5.13).

Media Sharing and Storage

Players store and share media related to gameplay as part of their player statistics. Video and images are the most common forms of media player dossier systems allow players to share. When a play analytic system allows players to share media, like videos, players are typically given the ability to create the media in the game. Call of Duty: Modern Warfare 3 and Driver: San Francisco allow players to capture screenshots/video during gameplay and upload it to a player’s personal dossier (REF see Figure 5.14). Players can then link to or send the files to other players, and players viewing the files can rate them for quality. Storing and sharing game content is described further in the Content Generation section in this chapter.

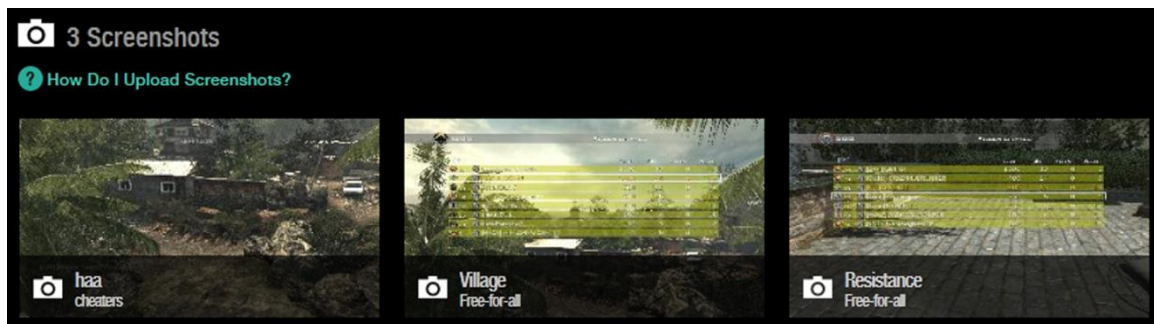


Figure 5.14: Elite allows players to share screenshots and videos captured from the game Modern Warfare 3.

Activity Feed

A continuous feed detailing events related to a player’s gameplay. Events within an activity feed include announcements related to the player’s statistics being captured. Earning achievements, gaining levels, moving up a rank on a leaderboard or recently played games are all examples of events that may appear within an activity feed. These feeds regularly appear on the main dossier pages for players and function as a sort of news or status update for players within the context of the game. Battlefield 3’s Battlelog,

Call of Duty's Elite and the World of Warcraft's Armory all offer activity feeds (REF see Figure 5.15).

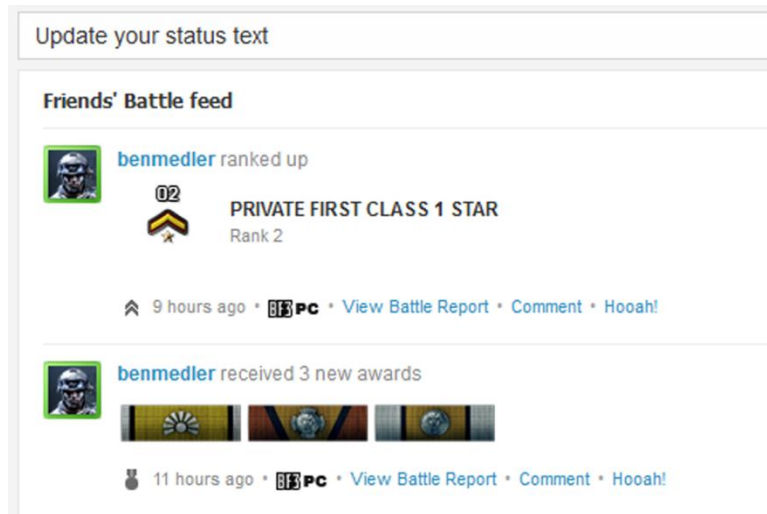


Figure 5.15: Activity feeds, such as the one in Battlelog, display historical events that are related to a player's gameplay or their friend's gameplay.

Friend List

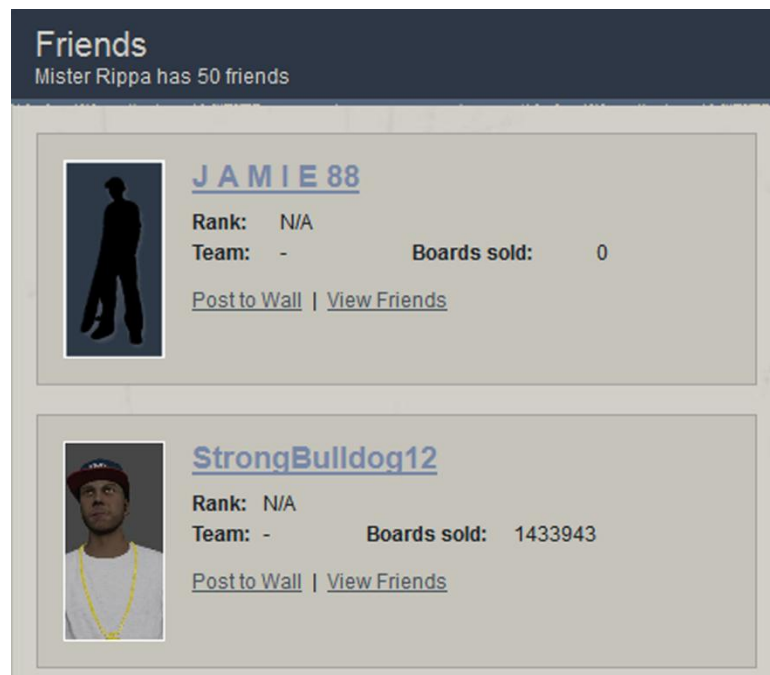


Figure 5.16: Some player stats allow players to create a list of friends consisting of other players who are in the play analytic system (Skate 3 pictured).

A list of fellow players that may be related to a player, i.e. a player's friends, or may wish to follow a player (REF see Figure 5.16). A player's friend list usually comprises other players known to the owner of the list and may be related to a player group (e.g. player clans). Some play analytic systems make it easier to share stats or data with friends versus other non-friend players. For example, on Battlelog when I add a player to my friend list their activity feed is added to my feed, making it easier to monitor their activities in the game. Other play analytic systems also allow a player to follow other players but are not 'friends' with the other players. A player following another player can monitor the followed player stats, but the followed player does not see any stats from the following player. Better performing players are often followed by other players and players who produce high quality game content may be followed as well (at least within a play analytic system with content generation capabilities).

Player Statistic Examples

Battlelog is the play analytic system built for the first-person shooter game Battlefield 3. The game is a modern military shooter with vehicle combat, both of which factor into the type of player stats Battlelog tracks. Player statistics are tracked for two types of game modes: co-op and multiplayer. Co-op gameplay, i.e. cooperative gameplay, revolves around completing missions based in Battlefield 3's campaign. Players can play each mission on different difficulty sections, which is recorded if they successfully complete the mission, and data from these missions (number of kills, vehicles destroyed and friends revived) is captured as accumulated statistics in the Battlelog system. Multiplayer matches in Battlefield 3 consist of two teams fighting one another over specific objects in the game. The objectives are usually physical areas in the game's environment where players must capture and/or maintain control over the objective locations for as long as possible. During each match players earn experience points for performing or accomplishing certain tasks. Killing an enemy player is one example task, but so is capturing a control point, spotting an enemy, reviving other players who are wounded or repairing a vehicle. Each of these experience point awarding actions are also typically captured and displayed as part of Battlelog's player stats. Some actions are accumulated over time, such as a player's kills or revive count, and other actions are

progressive, such as experience points which steadily progress a player through the game's leveling system.

Battlelog also has a number of other player statistic variations available in the system. There are a number of trophy rooms where a player's achievements, metals and awards are displayed. These trophies are typically awarded after each multiplayer match and reflect any exceptional gameplay a player exhibited. Earning the most experience points in the game is one type of award while earning the set number of points for defending a control point is another. Battlelog also allows players to compare one another based on any number of factors including their accumulated/progressive statistics and the trophies each player has won. Players can keep an eye on their friends as well since Battlelog has both a friend's list capability and an activity feed that announces events about both the player and their friends. Finally, Battlelog is not unique in the space of play analytic systems built for FPS games. Call of Duty: Elite has many of the features Battlelog offers in addition to other player statistic features like media sharing and storage, along with load outs of player equipment. Elite even offers the ability for players to create load outs within the Elite play analytic system and then upload them to their in-game persona to use in multiplayer matched. Brink and Bungie.net (when the systems were operational), and other play analytic systems built for FPS games, also provide a similar player statistic feature set to players.

Functional and Phenomenological Aspects of Player Statistics

Visible

Player Statistics is a category of play analytics that seeks to make personal play behavior visible. It is often empirical in nature. Gameplay events, or player actions, like finishing a level, killing an enemy or finding a hidden object are made visible through player stats because these events are tracked and compiled. Player statistics are basically an account of what happens to the player over time but only an account of what the system sees. If a particular event or player action is not monitored then it is not made visible by the system, and therefore is not represented as a player statistic. Some games

try to capture as much data as possible and to make as much of a player's gameplay visible. Systems like Battlelog, Elite, and the WoW Armory capture a lot of gameplay events from players, attempting to make multiple types of player behavior visible. But these example systems are part of very expensive, popular games. Many of the other player statistic systems do not provide the same level of gameplay event visibility as the larger systems.

Comparability

While other play analytic categories like leaderboards and match statistics are more suited for comparing players together, systems with player statistics sometimes offer the ability for players to compare themselves to other players. Usually, this comparison happens between friends. Having a friend list and the capability to compare players tend to go hand in hand within a play analytic system, the logic being if players are able to track their friends then they should be able to easily compare their stats. There are even cases where comparing players is part of the game. Assassin's Creed 2 used to have a web battle feature that allowed players to win awards for having better overall player stats than other players, which would often be players on their friend list. Player stats are often focused on giving players data related to their personal gameplay but there are cases when players can gain value from comparing their personal gameplay to other players.

Control

Player stats are always setup as a dossier versus a profile. Some player stat features allow players to add some content, maybe a character description or comment on a player's activity feed, but every player stat system automatically adds data related to a player's gameplay, whether they want the data added or not. Whoever controls the player stats therefore has an immense amount of control over what data is made visible and how the data is presented. And seeing as the only way to capture stats from players while they play is by having direct access to the game's source code, developers are often in control over player stats within a play analytic system. This means players rarely have the means to alter what data is collected and how it is displayed. Many systems have privacy

settings, players may opt-out of data tracking in some cases, but if the player wants to use the player stats features they have to use what the developer gives them.

The few cases where players gain control over their own stats are when a play analytic system offers an API or when players offer data to a third party service. Bungie.net and Spore offer APIs that link directly to their databases housing player data. Anyone who signs up to use an API can access this data and create their own play analytic system. Of course, the type of data captured is still controlled by the respective developer (a user has no control over how Maxis captures or organizes data recorded from Spore) but access to an API allows others to use the available data as they see fit. The second way players gain a little more control over their player stats is when they begin using third-party services. GamerDNA, for example, captures data related to what games a player plays and when players earn achievements. Since GamerDNA does not have direct access to data from the games the service tracks the data captured from players is not as detailed. However, the data Gamer DNA does collect is still useful because the service can recommend games to players based on the types of games that player has played in the past. The reason players have more control over their data is that each player can willingly become a member of these third party services. If a player does like how the service is using their data or how data is being presented, the player can leave. Being able to leave isn't always an option when using other developer produced play analytic systems like Battlelog or COD: Elite. Data is continuously captured from the player whether they like using those systems or not. At least on a third-party system, players can choose whether they want to continue to provide the service with data or if they wish to move to some other service.

Exposure

Almost every play analytic system with player dossiers are either public, meaning anyone can access a player's stats, or allow anyone with a system username/login to access a player's records, which is private to outsiders but easy to gain access. Exposure of player data is therefore rampant and typically quite easy to access. User profiles on social networks like Facebook and Google+ tend to be public as well, unless a user

manually blocks outsiders from accessing their accounts. This does happen with a play analytic system but to a lesser extent. Features like leaderboards routinely upload scores along with a player's gamertags (making it easy to find and track players across multiple games), though on rare occasions a game may ask the player whether a leaderboard score should be uploaded. Other games have been known to allow players to turn player data tracking on or off (e.g. Mass Effect 2 and Dead Space 2). But these examples are not in the norm for the same reason that social networking companies want users to have public profiles; public profiles means outsiders can see how many people are using the system and become enticed to join.

Player stats can be designed to be a centralized hub within a play analytic system; a hub where all other data related to a player is exposed. If we return to the concept of bracketing, the notion of determining the essential properties of an object while eliminating outside factors, player statistics is the play analytic category that brackets players within a system. The other play analytic categories are routinely linked with player stats too. Gamertags on a leaderboard lead back to each player's stats. Matches a player participated in, which lead to individual match stat pages, are often displayed alongside their player stats. Group stats and any content a player generates, if the game or play analytic system allows content generation, are linked to a player's dossier. This makes player stats the most personal category of play analytics, everything revolving around an individual player within the context of a game relates back to player stats.

Since player statistics are a method of bracketing player data within a game context, and play analytic systems tend to be publically viewable, the bracket around a player can cause an increased level of exposure. If, for example, a player has a public Steam account outside parties can monitor when the player is online, what games they have recently played or how often they have played in the last two weeks. This may not seem like an immediate problem but there have been reports of people being fired for posting comments about their job on blog or having racy photos on their social network profile (Solove, 2007). Employers could potentially use a system like Steam to track when their employees are playing games and whether they are playing while on the job.

This is an extreme example and many players use gamertags, not their real names, so finding someone within a play analytic system can be harder. However, so long as play analytic systems make player data publicly available there is a risk outsiders can misinterpret or misuse the data for other purposes. Both developers and players should be concerned about how play analytic data is exposed to players inside the system and to users outside the system.

Interconnected

Even though many player statistic systems collect and display data automatically, they also afford players many ways of connecting with each other. For instance, friend lists, comparing player features, and activity feeds allow players to monitor one another. Activity feeds in particular can be used to share high scores or events where a player earned an achievement. Battlelog for example allow players to comment on events that appear on a player's activity feed. This opens the door for players to talk about specific matches they played, share strategies they used or brag about their accomplishments. Another feature like media sharing also give players the chance to capture and reminisce about their past gameplay with other players. Bungie.net players often uploaded videos related to their gameplay (when it was operational), either because the video showed a specific player's excellent performances or something funny/impossible happened during a game (such as a vehicle getting stuck in the side of a mountain). Content like videos can often be rated and commented on by players too, creating further connections between players. There are also cases, in games that allow content generation, that players become popular as content creators. Games like Little Big Planet, Mod Nation Racer and Spore, all of which are games that thrive on player generated content, allow players to monitor the players who they feel are creating exceptional content in the game. Player statistics are therefore more than just a list of stats accumulated over time but affords players the ability to share their gameplay experiences with their friends and a community of players.

History

Player statistics are often ahistoric in a similar way leaderboards can be ahistoric. Most of the time the stats or data made available as part of a player stats (or dossier) page

consist of up-to-date data. Accumulated and progressive stats increase over time but offer no sense of how they have progressed over time. The trophy room variation I discussed in the section above is actually a fairly accurate description of player statistics over all, in regards to how player data is stored over time. Most player stats features can be seen as trophy rooms. The number of kills, matches won, levels completed or items collected are really trophies players present to others (not to mention all of the actual trophies, achievements, awards and metals that appear within a player's dossier). Of course players can share their data, discuss it with other players, and can monitor players over time but it is hard to keep track of how a player's data has changed over time.

The reason player stats are ahistoric may be due to how games in general are designed. Many games have a leveling or experience systems that increases over time as a player plays more and more. As players play they progress through levels and after reaching certain levels, are given access to more content in the game. Going back and looking at where a player has been is not as important as continuing forward and unlocking the next piece of content. This is a design prejudice many games have and it has filtered into how play analytic systems are setup too. There are examples of play analytic system displaying player stats temporally, often related to match statistics, but player stats are generally regarded as a way to represent a player's current status within a game instead of their history playing the game.

Player Statistics Summary

Player statistics provide players with way of tracking and exposing their gameplay progress. Many player stat features work like glorified trophy rooms where the achievements and events of a player's gameplay are put on display. Players can monitor each other using variations like activity feeds and friend's list, keeping up to date on each other's accomplishments and gameplay experiences. Variations like comparing players and rankings means players stats act as a source of competition between players too. Ranking players and comparing stats between players keeps them engaged competitively, in a similar way leaderboards can increase competitiveness through comparing scores. Additionally, developers often control what player stats are collected and displayed they

also control the identity of the player within their system. Systems like Battlelog, Bungie.net, COD: Elite, My Killzone and Social club present players stats automatically, as dossiers, within their respective communities. Players are represented by the dossier reports that the developers decide is relevant to the community. Developers also want as much player data to be public as possible because it helps expand the community. The more players exposing data the more it drives other players to take part and use the play analytic system. Player stats, thus, become a centralized hub for individual players where all other data related to a player is exposed and can be explored by players.

Group Statistics

Group statistics cover the collection and visualization of data from a group of players instead of individual player (i.e. player stats). Playing with other players, either cooperatively or competitively, is a critical aspect of many games. Battlefield 3 has a co-op (cooperative) ability for two players to simultaneously play the same level and earn points for achieving certain milestones. Games like Brink, League of Legends and Starcraft 2 place groups of players into competitive matches against each other. Players regularly form groups in order to socialize, tackle larger objectives in a game or compete competitively in tournaments. Playing with or against a group of players are common experiences in games and play analytic systems often reflect the desire of players to form groups by collecting and presenting data related to a group. However, how the data collected from a player group is presented to the group differs from system to system.

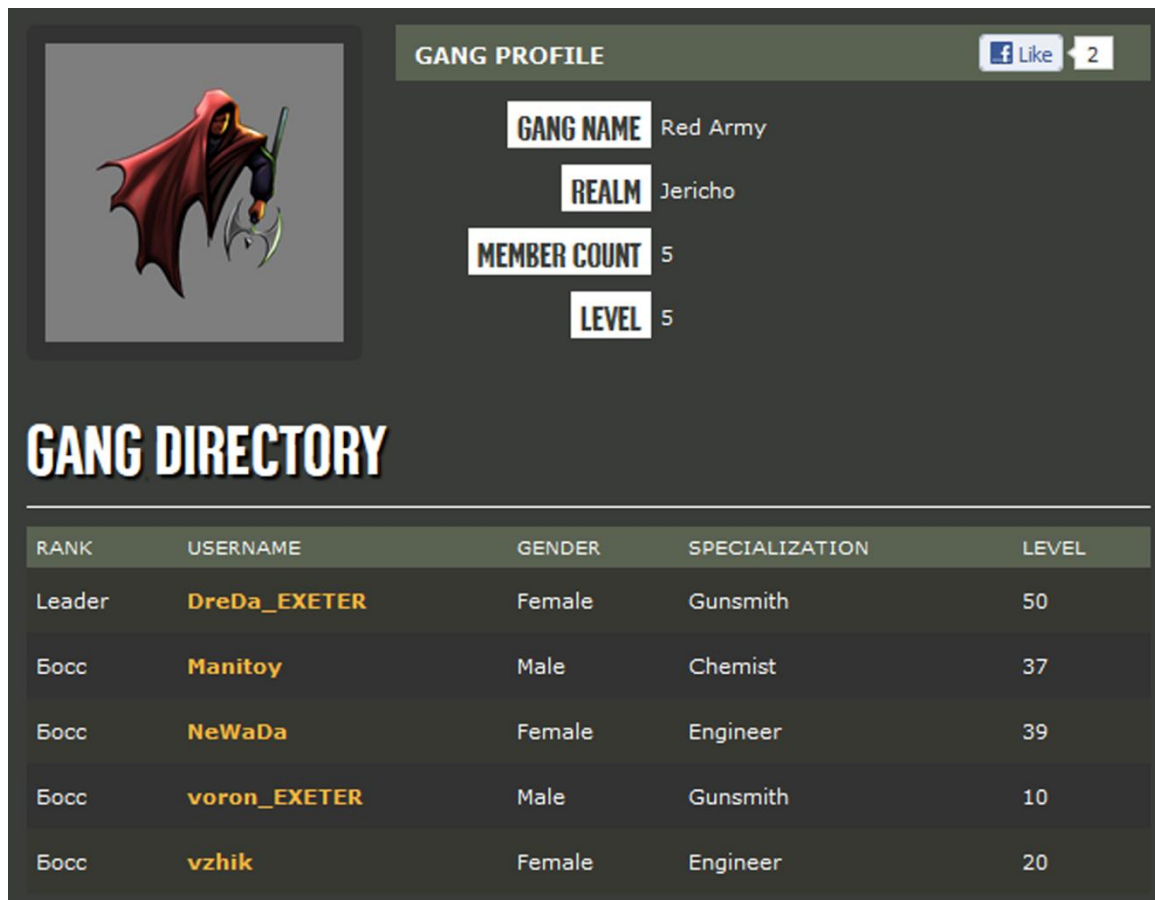
17 of the 81 play analytic systems offer some form of group statistics. Of course, players can form groups regarding any game they wish, there are many non-play analytic systems and services available to help connect players with a common interest in a game. What the 17 play analytic systems surveyed provide is a way of acknowledging a player group's existence and represents the group in the system. How a system acknowledges and represents a group, although, varies. Some groups are represented as a formality, where a play analytic system acknowledges the group but no real statistical data is presented regarding the group's gameplay. Other systems allow players to become fans of a real life sports team, or other type of common identity, and those fans form a group

around the team. Finally, the systems that allow players to form their own player groups, and collected data from the group, capture different levels of statistics based around the group. Sometimes a group is represented by adding together, or averaging, each individual player's statistics while other systems represent groups as if they are their own entity, complete with their own set of collected statistics. Even when groups are monitored and allowed to form within a play analytic system there experience as a group may be quite different.

Group Statistic Variations

The following design variations of group statistics appeared in the play analytic systems surveyed:

Acknowledgement Group



GANG PROFILE Like 2

GANG NAME Red Army

REALM Jericho

MEMBER COUNT 5

LEVEL 5

GANG DIRECTORY

RANK	USERNAME	GENDER	SPECIALIZATION	LEVEL
Leader	DreDa_EXETER	Female	Gunsmith	50
Bocc	Manitoy	Male	Chemist	37
Bocc	NeWaDa	Female	Engineer	39
Bocc	voron_EXETER	Male	Gunsmith	10
Bocc	vzhik	Female	Engineer	20

Figure 5.17: Not all group statistics are detailed. The acknowledgement group found in Crimecraft's group stats only provides a group's name, realm, level and member roster.

An acknowledgement group is merely a formality and doesn't affect how the play analytic system functions or what is presented to the group of players. Crimecraft and Lord of the Rings Online, for example, show almost no details regarding the groups within their system (REF see Figure 5.17). Players can join a group and there are pages representing the group but few statistics are gathered about the group.

Collection Group

Statistics from each individual group member are combined to form the group's overall stats. A collection group is a sum of its parts. Each player's stats are combined together, typically averaged, to describe the stats of the group. Steam groups display the total combined time the group's members have played various games on the Steam platform. Battlelog displays the best, average and worst value for stats related to character classes, kill/death ratios and score per minute (REF see Figure 5.18). Collection group stats are basically player statistics combined together.

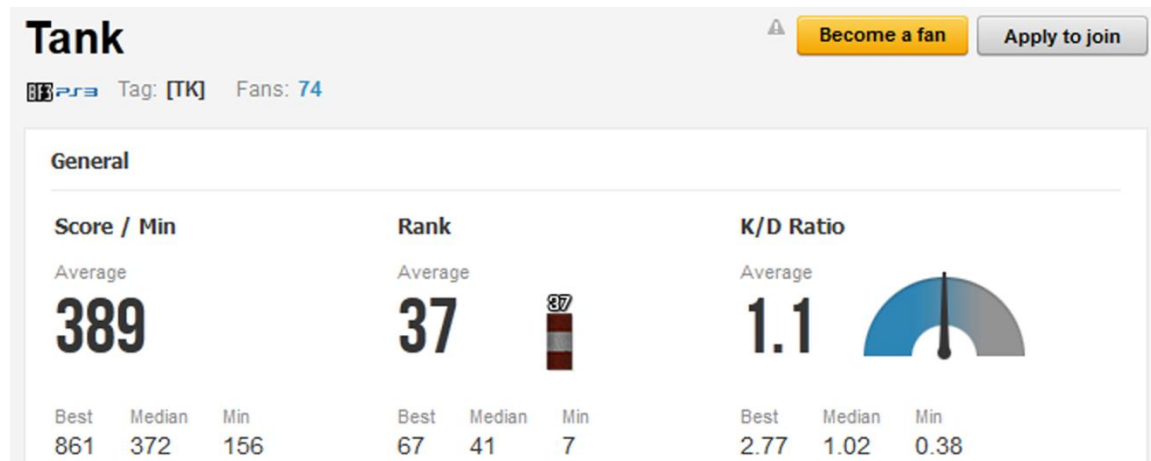


Figure 5.18: Battlelog groups are setup like collection groups. Each group member's stats are combined and averaged to form the overall group's stats.

Integrated Group

A group seen as a single entity capable of activities and accrues stats that individual players cannot achieve. Integrated groups are holistic entities compared to collection or acknowledgement groups. Each member contributes to the overall group stats and each group has stats directly related to the group itself. In some cases groups have stats related to tournaments, or other group related events, the group takes part in, as

the Brink, COD: Elite and My Killzone system provide. Some systems also give groups special awards or perks as their members gain experience points for the group while playing the game (REF see Figure 5.19). COD: Elite and WoW Armory link those awards and perks to the group's level. As players gather group experience points for completing achievements or events their group's level goes up. At different levels the group gains special abilities or features such as healing group members faster or extra experience points. Integrated groups gain abilities and stats individual players cannot receive on their own.

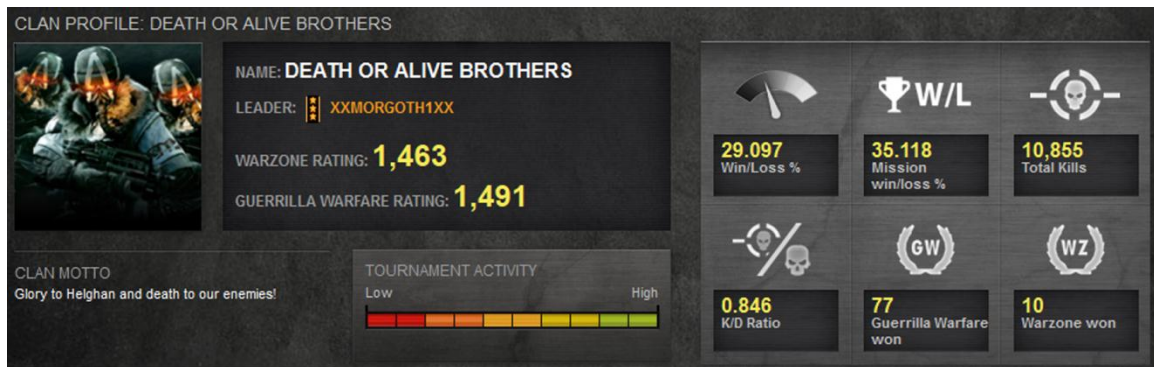


Figure 5.19: Integrated groups, like the groups found in Killzone 3, have their own stats apart from each group member's personal player stats.

Support Group

A group where players list themselves as supporting or fans of the group. Support groups represent groups players identify with but does not act as a group where the members regularly play together. For instance, EA Football Club allow players to become fans of real world football teams (REF see Figure 5.20). COD:Elite allows players to tag themselves in different groups. Players can make up tags or choose from a set of common tags (geographic location, military units, etc.). Each tag in Elite works like a collection group where each members stats are added together and there are personal group leaderboards players are ranked on. However, players can join multiple groups based on different tags. Support groups allow players to be a part of a group without having to commit to playing with the other members within the group.

CLUB DETAILS: VILLARREAL CLUB DE FÚTBOL "B" S.A.D.




FANS 470 (5 NEW)

Current Season: **99** Season PTS [View Table »](#)

League:  **LIGA BBVA** Rank: **1ST**

Overall: Global Rank: **1ST** **118,996** Total AVG XP [View Table »](#)

Country:  **SPAIN**

Figure 5.20: Support groups allow players to list themselves as a supporter or fan of a particular group. FIFA Football Club allows players to become fans of real life football (i.e. soccer) teams.

Group Statistic Examples

Every type of group statistics does have one thing in common, a roster. The purpose of a group is to give a group of players the means to acknowledge themselves as a singular group. Rosters are the way a system acknowledges everyone in the group. Even acknowledgement and support groups have rosters, although both types of groups usually have little else in terms of group statistics. Generally, a roster is more than a list of names though, often each member is listed with a number of statistics. Values like a player's character or class type, experience points earned, kills made or matches won are listed next to their name. Rosters can also double as leaderboards using any of those additional stats to rank the players within the roster.

Besides rosters, groups have other special features such as combined statistics, match history and organization tools. Collected groups for example are groups where each individual member's stats are combined to represent the stats of the group as a whole. Battlelog combines each member's stats together while providing information on who has the highest value within each stat value (kills, assists, revives, etc.) and what range of values (highest to lowest) exist within each stat value. Activity feeds can work in a similar way as well, where each player's activity stream is combined to form a group activity feed (as seen in the World of Warcraft Armory). Matches may also be recorded

as group matches if enough members of a group take part in a specific match. In this case, instead of matches only showing up as part of a player's dossier, matches can appear on a group's dossier. Brink, for instance, had match histories for both players and groups. Finally, some play analytic systems include organization tools for groups to setup common times group members can play. Tools like calendars and invite requests allow groups to determine when a critical mass of members can be online and play together. Play analytic systems for MMOs (WoW and Lord of the Rings Online) typically have some sort of calendar feature because certain types of content in MMOs can take many players to complete and a certain level of organization is required. Invite requests are also used in games to invite an ad-hoc group of players to play. For example, Gears of War, which does not have group stats, has an invite system tied into its play analytic system, War Journal.

When a game has the integrated group variation as part of their play analytics system, player groups are tracked over time as a cohesive unit who additionally earn special awards as a group. For example, a group in World of Warcraft is known as a guild and each guild earns their own achievements and experience. By completing quests, crafting items, taking part in dungeon raids, and killing enemies each guild member contributes to the guild's achievements and experience. As guilds gain experience they progress along a series of level. At each level the guild, as a whole, gains new perks related to gameplay such as guild members earning more gold when killing enemies, increased storage for guild items or decreases travel time for members using special items. A similar perk system exists in COD: Elite too. As clan members, clan being the name Elite gives groups, take part in clan operations (weekly events where clans attempt to compete different objectives) they gain clan experience and gain levels. As a clan's level goes up, they too earn perks in a similar way guilds do as shown in the Armory system. Whether providing a group roster, combining statistics, allowing members to organize or unlocking perks over time, group statistics are about presenting a group of players as having related interests and goals within a play analytic system.

Functional and Phenomenological Aspects of Group Statistics

Visible

Group statistics are a way of making groups visible within a play analytic system. Every type of group variation is meant to acknowledge a player group's existence. Besides acknowledgement and support groups, some of the group variations additionally make other data visible. Collection groups make each group member's individual stats visible to the group as an ensemble. Sometimes these collection groups work like friend leaderboards, where each member's stats are compared and ranked against the rest of the members. Integrated groups present groups as a unique entity, with separate stats compared to each individual member. Data about how the group levels up and performs in competitive multiplayer are often made visible through integrated groups.

Comparability

Acknowledgement and support groups do not provide much capability for comparing data but collection and integrated group do. One way collection and integrated groups compare data is by comparing players to other players. Collection groups, especially, compare group members together because a collection groups' stat are the sum of the stats from each individual player. Players within a collection group can be compared in a similar way a friend leaderboard works, based on criteria like score, kills or wins. Groups can also be compared against other groups. Integrated groups, which have stats related to the group as a single entity, are often compared against each other, particularly when a game has competitive multiplayer. Systems for Brink, Killzone and Starcraft 2 all compare groups based on their performances in various types of multiplayer matches.

Control

Developers have most of the control over how a group is defined within a play analytic system and how that system tracks the group's gameplay. Players in a Battlefield 3 group only have the ability to form collection groups, My Killzone groups always have tournament statistics (even if a group of players do not participate in any tournaments) and World of Warcraft guilds can only unlock guild levels if their members help earn experience for the guild. The one thing developers do not have control over is what group

a player joins. Players can join any group they want and players within the group do not necessarily have to use the group features provided. Killzone players do not necessarily have to play in tournaments. Two players can start an EA sports football franchise (their equivalent of a group) and keep their group just between themselves. In most cases, group statistics are a bonus for players because players form groups regardless of group statistics a play analytic system provides.

Exposure

One thing groups provide players is a way to bracket their data within the context of other players they know or share interests with, those players within their group. In a similar way a friend lists is setup, members of a group expose their data to friendly players because most players joining a group do so in order to play with the other players in the group. Players want their group members to know their stats since they are the players who likely care the most. Plus, exposing data within a group can increase competition between the group members, much like how friend leaderboards work. Battlelog groups mark which players have the most kills or highest score per character class. Other players in the group see that and may try to take the top spot away from their fellow group member. Competitive play between groups also factors into how integrated groups operate. Group stats from integrated groups are exposed to the community at large and systems like Battle.net and My Killzone expose those group stats as part of competitive tournament or ladder features.

Interconnected

Forming a group is an easy way of saying the group members are interconnecting with each other. When a player is part of a group, and associates the group as part of their identity, they gain a connection to the other members in the group. While leaderboards or match statistics may connect players based on gameplay (whether through a set performance criteria or a specific match played) a group is a declaration of a constant connection between players and forms a cohesive identity for those players (or at least a more cohesive identity than say a friends list). Forming a group also allows the group to interconnect with other groups. Having a group means gaining access to features only

given to player groups and interacting with other groups using those features. Clans (i.e. a group) in the Darkfall Political map system can declare other groups their allies or enemies, for instance. Having a group therefore helps connect players within the group and connects them with other groups, often as a means of competing or cooperating with each other.

History

Much like player statistics, group statistics have little historic data. Acknowledgement and support groups barely keep track of any data at all, let alone present the data over time. Collective groups rely on stats from individual players and if a system only presents up-to-date player stats, instead of stats over time, then the group stats will not be historically focused. Only the integrated groups, where groups are seen as separate entities are their signs of systems presenting data historically. Some systems, like the WoW Armory and Darkfall Political Maps, keep track of events over time. Darkfall Political Maps particularly keeps a historic record of clan events related to: capturing towns, forging alliances with other clans and declaring war on other clans (player stats also keep track of when players join or leave a clan). However, the main feature of the Darkfall system is to produce a map of the current political standings in the game; nothing is done with the past event list besides allowing others to access the data. Group stats, therefore, share the same prejudices as player stats, data is presented as ahistoric because players are thought to be only interested in the current state of their game.

Group Statistics Summary

Players regularly form groups for many released games and those groups represent a part of their identity within a game. Group statistics help to solidify that identity by offering players the means to track their data in relation to the group. Play analytic systems with integrated groups recognize a player group as a cohesive entity, providing the group with its own set of unique stats. These integrated groups often appear in play analytic systems built for competitive games where a group of players compete against other groups of players. The Armory, Battle.net, My Killzone, and COD: Elite

each have competitive systems that track and rank groups of players in addition to tracking and ranking individual players. Other group types like acknowledgement, collective and support also offer players a way to represent themselves as a group of players within a play analytic system. Support groups allow players to rally around common themes like sports teams or geographic locations, while collective groups combine and compare each member's individual stats. Group statistics in a play analytic system allow players to compare, cooperate and compete amongst a smaller set of players (their group) while also offering the ability to compare, cooperate and compete with other groups of players too.

Match Statistics

Continuing on with another statistics category, match statistics cover the data gathered during a particular play session, otherwise known as a match. A match typically refers to a situation where a player, or players, compete against a common enemy for a finite amount of time; the enemy being other players or a computer controlled opponent. Sporting events like a baseball game or a football game are examples of matches. A single Starcraft 2 battle between two opposing teams is a match, as is a single battle in Halo: Reach. However, match statistics can refer to a play session, such as one that takes place during a single player campaign. Splinter Cell: Conviction keeps track of when a player began a certain level and who they played the level with, if the player was cooperating with another player. Matches, therefore, do not always mean a head-to-head competitive situation. Matches refer to repeatable portions of a game where stats can be collected during the finite amount of time the match is occurring (a chess match between two players for example).

Of the systems covered 17 of the 81 provide match stats. Many of the play analytic systems with match stats are for competitive games. As a result, the match stat variations found generally revolve around presenting each match as a separate report. Player and group stats can be viewed as a collection of a player's or group's gameplay over their entire gameplay experience. Instead, match stats cover a finite amount of that gameplay experience and are organized like individual reports covering the results of

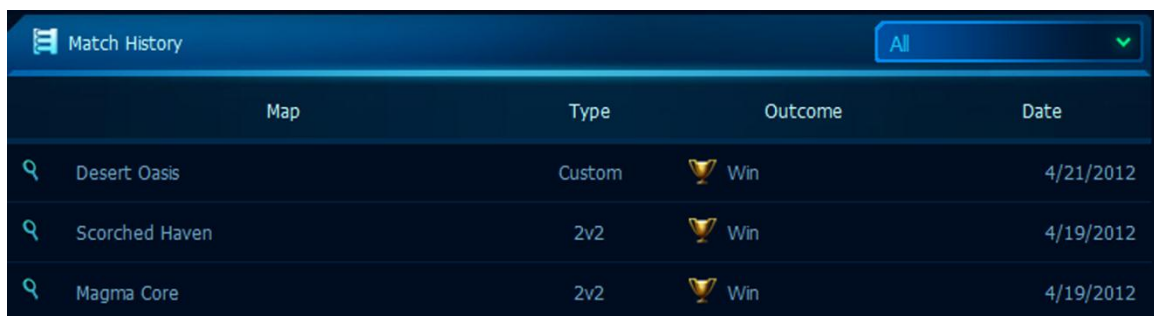
each match. These individual reports can be severely limited. The match reports in Crimecraft give the name of the mission (i.e. match), the time of the mission and the level the player ended the mission with, with no other data provided. But other examples, particularly systems like Bungie.net or SC2Gears, provide detailed reports covering moment to moment events that occurred in a match over time.

Match Statistic Variations

The following design variations of match statistics appeared in the play analytic systems surveyed:

Accumulated Report

Simple data about each individual match is accumulated to give an overview of the match results. Generally, this consists of stats regarding how many matches a player won or lost but may include: total time played, most recent match played or type of match played. Starcraft 2 matches displayed on Battle.net website display the map, type of match, whether the player won and what time the match was fought, without any further details regarding each match (REF see Figure 5.21). Otherwise, an accumulated report may accompany a list of recent matches a player participated in that lead to more detailed reports for each match.



The screenshot shows the 'Match History' window from the Starcraft 2 Battle.net interface. It features a dark blue header with a menu icon and the title 'Match History'. On the right, there is a dropdown menu currently set to 'All' with a green checkmark. Below the header is a table with four columns: 'Map', 'Type', 'Outcome', and 'Date'. The table contains three rows of match data, each preceded by a magnifying glass icon.

	Map	Type	Outcome	Date
🔍	Desert Oasis	Custom	🏆 Win	4/21/2012
🔍	Scorched Haven	2v2	🏆 Win	4/19/2012
🔍	Magma Core	2v2	🏆 Win	4/19/2012

Figure 5.21: Accumulated match reports provide players with overview stats from the events that occurred during the match (Starcraft 2 pictured).

Detailed Report

Contains extended data related to an individual match. Instead of simply saying who won the match, a detail report may break down a match based on whom killed

whom, what objectives were completed, what awards were given out, how points were awarded over time, etc. Match data from Halo: ODS, on Bungie.net, included details about how a match transpired down to individual waves of enemies sent against a player and their team (REF see Figure 5.22). A kill event is further broken down into what weapon was used, was it a headshot and what enemy was kill. Detailed match reports describe data about a match instead of describing the results of a match, as an accumulated report does.



Figure 5.22: Match stats that offer detailed reports give an extensive look at the events that occurred during a match (Bungie.net pictured).

Event Timeline

Temporal attributes are given to events that occurred in the match and presented over time. SC2Gears, for example, provide timelines detailing when players built units, mined resources or attacked their opponents (REF see Figure 5.23). First-person shooter play analytic systems (COD:Elite, Gotham City Imposters) present connected kill and death events over time and the system for Gotham City Imposters allows a player to animate these events like a replay.

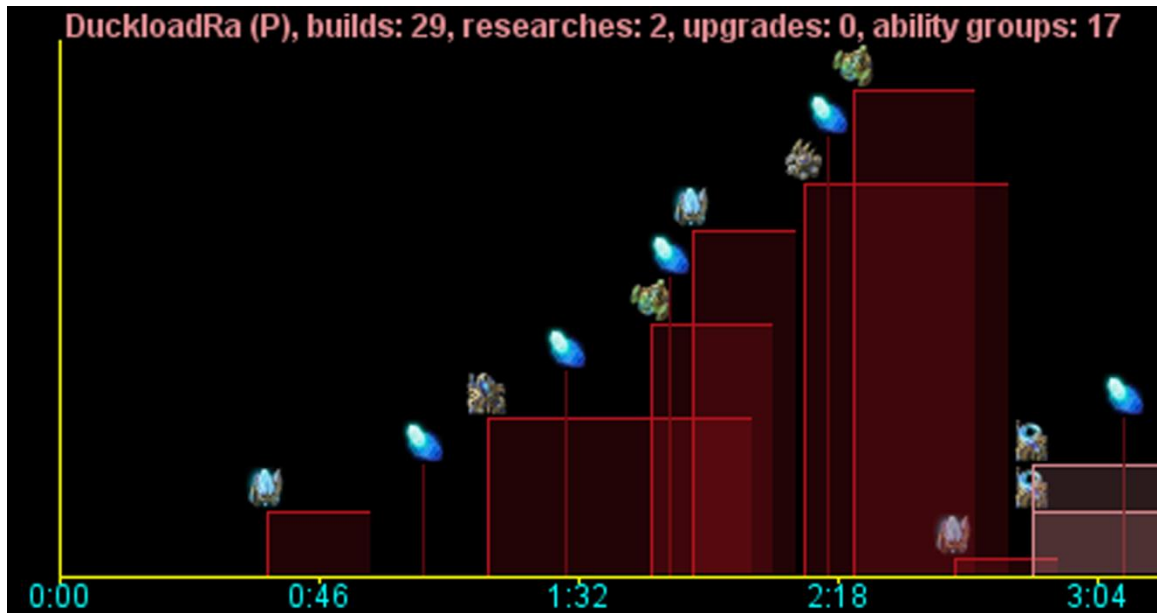


Figure 5.23: Timelines are used in match stats present data temporally based on when certain events in the match occurred (SC2Gears pictured).

Match Leaderboard








US ATTACKER					
#	Sq	Soldier name	K	D	Score
1	B	 HpxT4nk	28	10	4207
2	E	 InsIJrgent	23	10	3200
3	B	 FOMIN77	7	10	2942
4	A	 McWells	15	6	2133
5	C	 delete180	11	11	1609
6	C	 Knaledge	11	6	1498
7	A	 TwoHandGuns	9	9	1296

Figure 5.24: Individual matches can have their own leaderboards and players who participated in the match are ranked by their individual stats collected during the match (Battlelog pictured).

Players are ranked within a match based on certain criteria. Usually refers to how awards are handed out based on match results. Players may be ranked by most kills, experience points awarded, accuracy, etc. Players may also be ranked within their team

instead of amongst everyone in the match (REF see Figure 5.24). (Also see the Leaderboard section in this chapter for further explanation of leaderboards in general).

Match Statistic Examples

The range of the data made available to players in a play analytic system after they have completed a match varies greatly from system to system. Brink, Crimecraft, and Splinter Cell: Conviction and Starcraft 2 (Battle.net) provide limited match statistics to players, often consisting of terse variables like kills, scores or who won the match. These systems exist even though some of the games (Starcraft 2 being one of them) offer a more detailed report of a match as part of the in-game summary report that appears after a match concludes.

As match statistics become more detailed, the systems that provide the match stats focus on different aspects of a player's gameplay. Halo Wars, for example, has many accumulated stats associated with collecting resources, using abilities and structures built. These data points are collected because Halo Wars is a real-time strategy game where collecting resource, using faction abilities and building structures are major parts of the gameplay. Producing and controlling units is another major part of Halo Wars and, in addition to the other stats provided, a detailed report covering what units were produced, when they leveled up, what units they destroyed and what enemy units destroy them are laid out for each player too.

Offering more data about a match makes each match's data more valuable for every participating player. Match stats give players extra recognition for their gameplay, beyond other player statistics collected for each player. Halo: ODST's single player and co-op mode, for instance, have timelines plotting a player's score over time during a match and when specific waves of enemies attack. Other scores and variables in the game are broken down for each player, including: what enemies they killed (broken up across all the enemy types), how many kills they earned with each weapon (broken up across each weapon), and what metals were awarded during the match. All of these stats can be

compared amongst the group and the match acts as a historical record of how player stats are accumulated across time.

Other players can benefit from reviewing match stats because the matches offer a peak into how the participating players in the match approached playing the game. Using a system like SC2Gears, players can review Starcraft 2 replays in striking detail (Starcraft 2 being another example of a real-time strategy game like Halo Wars). Every major game event is visualized and in most cases is projected across time (i.e. the length of the match). Someone reviewing a match can find out how many actions per minute a player achieved over time, when and where buildings were placed, when units were produced, where groups of units were moved, when battles took place and how many resources a player had over time. Analyzing these reviews can help players discover strategies or best practices they can apply to their own gameplay. But discovering strategies can only occur if a system like SC2Gears offers enough data for other players to review, unlike the terser systems we see for Brink or even Starcraft 2 (battle.net).

Functional and Phenomenological Aspects of Match Statistics

Visible

Match statistics make a certain period of gameplay time visible. Unlike player and group statistics, match statistic features reveals a particular time period of gameplay and treats it as an entity (in a similar way player stats treat a player entity, or groups stats a group). Since match stats are finite – they represent a particular moment in a player's gameplay history – the data made visible is always a type of historical record. Even when only accumulated stats are collected, the fact that multiple matches are collected over time means they provide a greater ability to see how players evolve over time. However, almost no play analytic systems present and analyze matches over time as an analysis of a player's evolution. In terms of type of stats found in match stats, the values found in the player or group stat features visible. If a player's stat dossier records how many kills a player has made overall, each individual match will most likely state how many kills the player earned. Additionally, detailed reports within match stats may present moment to

moment events related to such values, like kills, and makes match data visible enough to almost create a replay of the match.

Comparability

Comparing between matches is rather limited in almost all play analytic systems with match stats. Matches are typically presented in a temporal list, where the latest match is shown first. These match list may be made to look like multiple criteria leaderboards, where various accumulated values about the matches are compared (e.g. the winner of the match, how many kills were made, how much experience points were earned), but do not necessarily rank the matches based on these values. Only SC2Gears offers in-depth analysis of multiple replay files at once (replays being highly detailed match data files). The system can compare both matches and players, even giving reports about a player's development over time based on the provided matches.

Comparing players within an individual match is more common than comparing whole matches to one another. Players may be compared using a leaderboard where different values earned during the match are used to rank each player. Values like experience points earned and total kills are common values used to compare players. Other systems like Bungie.net, COD: Elite and the system made for Gotham City Imposters also track player values and events over time. Those systems are able to show when players were killed or objectives were completed along a match timeline. These values are often presented using a map of the level, making it easier to compare spatially where the events occurred.

Control

Like other game statistic categories developers usually have control over what data is collected from a match. Player and group statistics are often calculated based on events collected from matches. If a developer controls the data collected during a match then they control what data is available for other statistics (whether player, group or global stats). Developers do not have control, however, over what happens in a match and this has allowed players to step in and take control of their match data.

Replays are files filled with enough data about a match to recreate the match for later viewing. “Enough data” usually means a replay has more match data than is officially provided to players. Match stats in a play analytics system may present the player with how many kills they earned but a replay has a record of when the player made a kill, along with everything else happening during the match at the time of each kill. Therefore, replays have been a source of valuable data players can use to analyze their matches. SC2Gears, for instance, is a player produced replay analyzer for Starcraft 2. SC2Gears provides a detailed analysis of Starcraft 2 replays, more so than anything released from the developers of Starcraft 2. Players are able to review most of their data over time instead of as accumulated statistics, and even compare whole matches to one another. SC2Gears is able to provide a detailed analysis of Starcraft 2 matches because Starcraft 2 replays contain much more data than is released as part of the play analytic system maintained by the developer. In order to get around the developer made systems, players build parsers to parse the replay files in order to access the extra data not being made visible. Although, at times parsing replay files can be difficult because developers may use some proprietary storage method (such as Blizzards (MPQ, 2012)) Control over replay data, and match data, can therefore come down to whether players take the time to reverse engineer replay files.

Exposure

Match statistics are the only type of statistics that have the ability to accurately expose how a player behaves in a game, when referring to the quantifiable actions players perform in a game. Detailed match reports visualize captured player actions and physical movement across a finite amount of time. Compared to player statistics, which can show a player’s tendencies based on accumulated values, match statistics use that finite amount of time to bracket a player’s gameplay data. While match stats can contain accumulated stats (e.g. total kills, cars passed, experience points awarded) the fact that the accumulated stats are representative of a smaller window of time makes the data much more specific to a player’s gameplay than accumulated stats collected over the player’s entire experience with a game.

Matches also tend to capture an increased amount of data during the bracketed time frame representing the start and end of the match. The increase amount of data can provide more context to the player's behavior when compared to other accumulated stats. An increase in context, based on an increase in available data, means different levels of analysis can be done with match stats (for example, SC2Gears can plot a player's performance between matches, which would otherwise be impossible with accumulated player stats) but also means more player data is exposed publically. Accumulated player stats can be argued as relatively more private compared to match data because match data allows a way to voyeuristically monitor a player's gameplay. Replays, for example, are distributed through a few of the systems surveyed (e.g. SC2Replays.com). Spectator players can access these replays and re-experience what the recorded player(s) experienced. Every in-game occurrence can be witnessed by outside, spectator players, which is not possible when exposing other player, group or global stats. While match stats provide an accurate depiction of a player's gameplay behavior they are the most exposed type of play analytic stats.

Interconnected

If match stats can expose a player by releasing actual gameplay actions than match stat can also expose the interconnections between players within a match. Single player matches are rarely recorded by play analytic systems (although games like Starcraft2 do allow players to record single player replays) and there tends to be multiple players recorded within each match. Match stats, therefore, afford the ability for players to realize the connections between their gameplay and their opponent's gameplay. These connections can be realized after the match concludes and players use the match stats to reflect or reminisce about the past match. Reflection after a match concludes can be beneficial because while a match is commencing a player's focus is on their own immediate situation and the data they have available to them at that time. Match stats and replays give players the chance to return to a match after gameplay is over and reflect on what happened in the game. For one thing, players can determine what other players were doing during key moments of the match. For example, in a Starcraft 2 or Heroes of

Newerth replay players can track other players over time, determine what strategies those players used and when did the paths of certain players cross. Players can also reminisce about interesting moments that occurred in the game, such as when a good move was made or a funny moment occurred (I have defined these moments as exquisite gameplay (Medler, 2009b) and has been defined as impressive play (Lowood, 2010)). The interconnections players experience through match stats is both reflecting on how their gameplay intertwined with their allies and opponents, while also reminiscing about interesting events that occurred during gameplay.

History

Match stats are the only type of gameplay stats that have an automatic sense of history. Even if a play analytic system only offers accumulated values after a match is completed, matches represent a finite moment in time and can be compared with one another across time. Although, there are a number of play analytic systems that provide detailed match stat beyond accumulated stats. Bungie.net. COD: Elite, Gotham City Imposters and SC2Gears all use visualizations to chart certain events that occurred in a match over time. Gotham City Imposters and SC2Gears even have playback features that incrementally reveal those match events so players can see when, and where, events occurred in their natural temporal order. Providing such features does make sense; if a play analytic system is recording more data within a match why not build features that can make use of the increased amount of data?

Even though match stats reveal data temporally, and often in greater detailed than other stat features, this does not eliminate the prejudices the systems have for displaying match data. Most of the developer built systems that spatially visualize matches over time tend to focus on a limited number of events. Gotham City Imposters, for example, only shows kills and deaths on a map, while COD:Elite shows kills, deaths and objectives completed. These systems do not track where players help another player (players may heal or revive other players, if the ability is in the game) and typically do not show the position of players over time. Match stats also share in the same type of prejudices player stats have. If a stat is recorded during a match it will most likely show up, in some form,

in a player's stats, but if certain stats are not recorded than the actions behind those stats remain hidden. A game's developer therefore still controls what data is captured during a match and the decisions made to capture specific stats are the prejudices a developer places on a play analytic system.

Match Statistics Summary

Matches represent smaller chunks of time throughout a player's entire gameplay experience with a single game. A Starcraft 2 player may play hundreds, or thousands, of matches over the course of their time spent playing the game. Play analytic systems built for games that structure the gameplay experience into a series of matches frequently collect statistics related to each individual match. Match stats are typically more detailed than other forms of stats. For instance, Gotham City Imposters visually maps where, and when, each player died and earned a kill during a match but only adds the total number of times a player died and earned kill to a player's personal player stats. Time and space are often dimensions represented as part of match stats even though these dimensions do not come up as often in other forms of statistics. Matches can give a better sense of a player's behavior because details like temporality and spatial position are combined with gameplay events. Systems like SC2Gears allow players to recreate and analyze any Starcraft 2 match, visualizing the data over time and spatially on the matches map. This means that player data is the most exposed when it comes to match stats but players also have the greatest capability to analyze their gameplay behavior.

Global Statistics

Unlike the other types of play analytic statistics, global statistics combines a game's community as a holistic entity. Instead of describing data related to a single player, a single group of players or a single match, data collected from every player is added together and presented as a whole. Sometimes global stats are as simple as saying how many players are online. Systems like EA Sports Hub has a running total of the players online at any time on their home page and Steam presents a graph in their 'Stats' section of their system that plots the number of concurrent Steam users throughout each day. Other examples of global stats treat player data as smaller pieces of a larger effort.

Variations like community events or global war features take each individual player's data and turn it into a combined effort by all players towards a common goal. Whether that goal is to unlock content or defeat an enemy faction, global stats attempt to make the entire player population's data relevant to all players.

Of the play analytic systems surveyed, 20 of them had some form of global statistics. This is slightly more than the systems with group or match statistics. The reason why global stats are as popular compared to group or match stats, even though many players form groups around most games and many games are based around playing matches, may be due to the fact that global stats are a way for game developers to promote, even brag, about their player community. When Bungie.net or FIFA release data saying millions or billions of matches have been played in their respective game franchises that is as much a marketing plug as it is a piece of data covering a global community of players. Of course, every piece of data released from a commercial game company can be argued as being related to marketing or promotion; using statistics to keep players engaged with a game's content is one of the arguments of this dissertation. Global stats are the type of stats that are easier for companies to give out because they are often extremely large numbers, which promote the game, while being too general for other parties, like other game companies, to take advantage of. The only number that rarely is given out in a global stat system is the total number of players in the system, which could be informative to other parties. Global stats are the only type of stats that are both beneficial for the game company, as a promotional tool describing the player community, and beneficial to the players, as an acknowledgement of the player community, on equal ground.

Global Statistic Variations

The following design variations of global statistics appeared in the play analytic systems surveyed:

Accumulated Statistics

Common events or actions players perform in a game are collected and accumulated into a set of statistics covering all players. This means any player contributing data can add to the accumulated stats. Examples include total values of: play time (often in counted in years), number of kills or other specific events, and number of content files available (levels, characters, missions, etc.). If accumulated statistics appear on a webpage they often are positioned on a hub webpage (REF see Figure 5.25). These hub webpages are at the top of the webpage navigation hierarchy, are often the first page players see (most likely because they have to login), and lead players to their personal stats or other player, group or match stats.

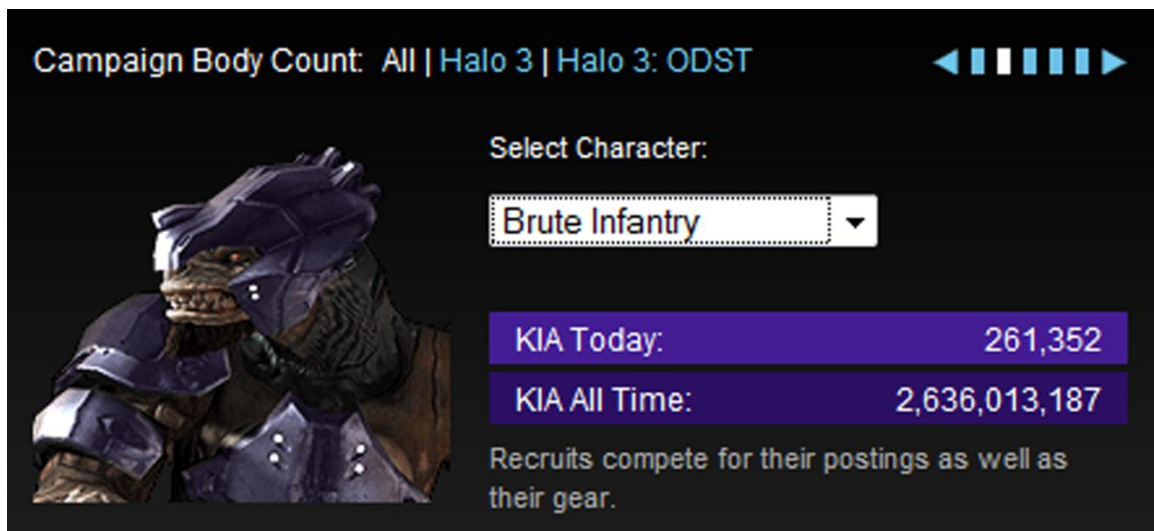


Figure 5.25: Bungie.net displayed a number of global stats consisting of all the current Halo 3 and Halo: ODST games played over different time periods.

Community Collection Event

A collection quota related to a community wide event is set and the game's community of players, as a whole, must perform a certain amount of actions in order to meet the quota. A community collection event could stipulate that 100,000 games need to be played before a new map or level is released to the community, for example. Community collections are similar to accumulated statistics because they collect particular actions or events from an entire player community. Although, community collections are unlike accumulated stats because they have a specific purpose (one

regularly attached to the release of extra game content) and have an eventual end point where further actions are not accumulated. World of Warcraft had a community event collection in 2006 called the ‘Ahn’Qiraj War Effort’ where players on each WoW server had to collect thousands of in-game items before a new area of the game would be available on each server (REF see Figure 5.26). After the collection was completed on each server the area was hence forth opened to all the players, without having to continue to accumulate more items.

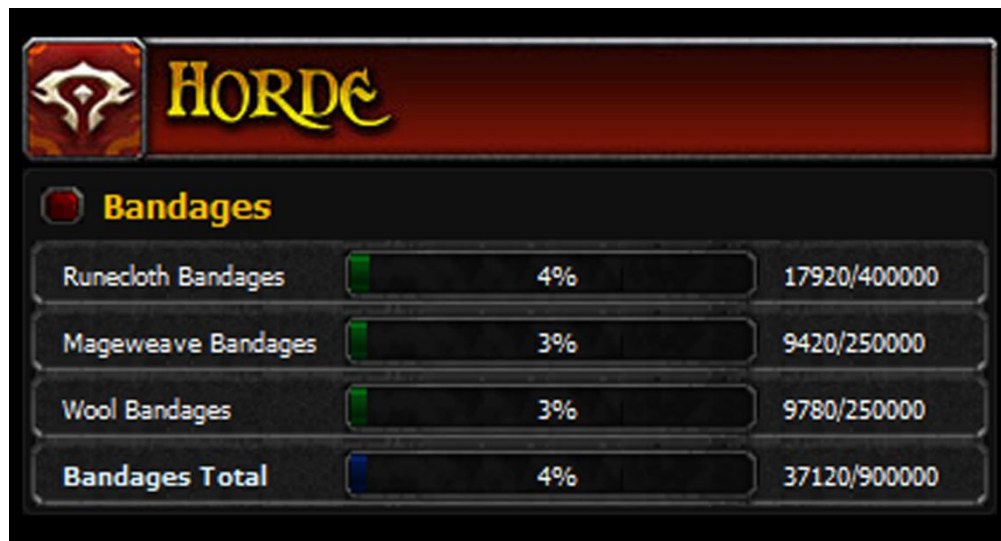


Figure 5.26: During the ‘Ahn’Qiraj War Effort’, a community event in World of Warcraft, players had to submit resources in the game and a system kept track of the resource collection progression.

Global War

Tied to battle or fighting-based games, global war systems accumulate results from individual game matches and use them to define which game factions are winning the overall “war”. Just like the other global statistic variations, events are being collected from all players and accumulated into total values, in this case match wins are being collected (REF see Figure 5.27). Brink, for instance, has two factions (Security and Resistance) in the game and players play as either faction when playing a match. If one faction wins a match, let us say the Resistance faction won, that “win” contributes to the overall “war” where the Resistance now has one more win compared to the Security faction. Whichever faction has the most wins is said to be winning the war (other values like kills or objectives completed during a match can be used too). Usually a global war

system has many sections or areas for players to fight over. Sections are often tied to specific maps or levels where individual game matches take place. At any point in time, one faction may have more wins in one section but have fewer wins in another. This simulates a war where each factions is attempting to take over sections of a war torn region, and each game match is used to simulate a single battle taking place as part of the war. How many “wins” each faction has changes over time as more matches are fought and as older data is taken out of the system. For example, a global war system may only accumulate data for a week’s worth of time before removing the data that is over a week old.



Figure 5.27: The game End War has an additional ‘Global War’ feature that took the results from player matches and aggregated them together into mimic a larger war effort.

Usage Statistics

Specific in-game characters, items or levels are tracked across all games to see how often players use each piece of content. Usage statistics are similar to accumulated statistics but relates to specific content within a game, as opposed to actions or events such as kills or wins. Heroes of Newerth tracks how often each hero character and items are used in every match (REF see Figure 5.28). Each player chosen a single hero

character and can buy multiple items in each match. The usage statistics show how popular heroes and how many players buy specific items each day (or some other specific time period). These stats are then shown over time and can be compared against each other.

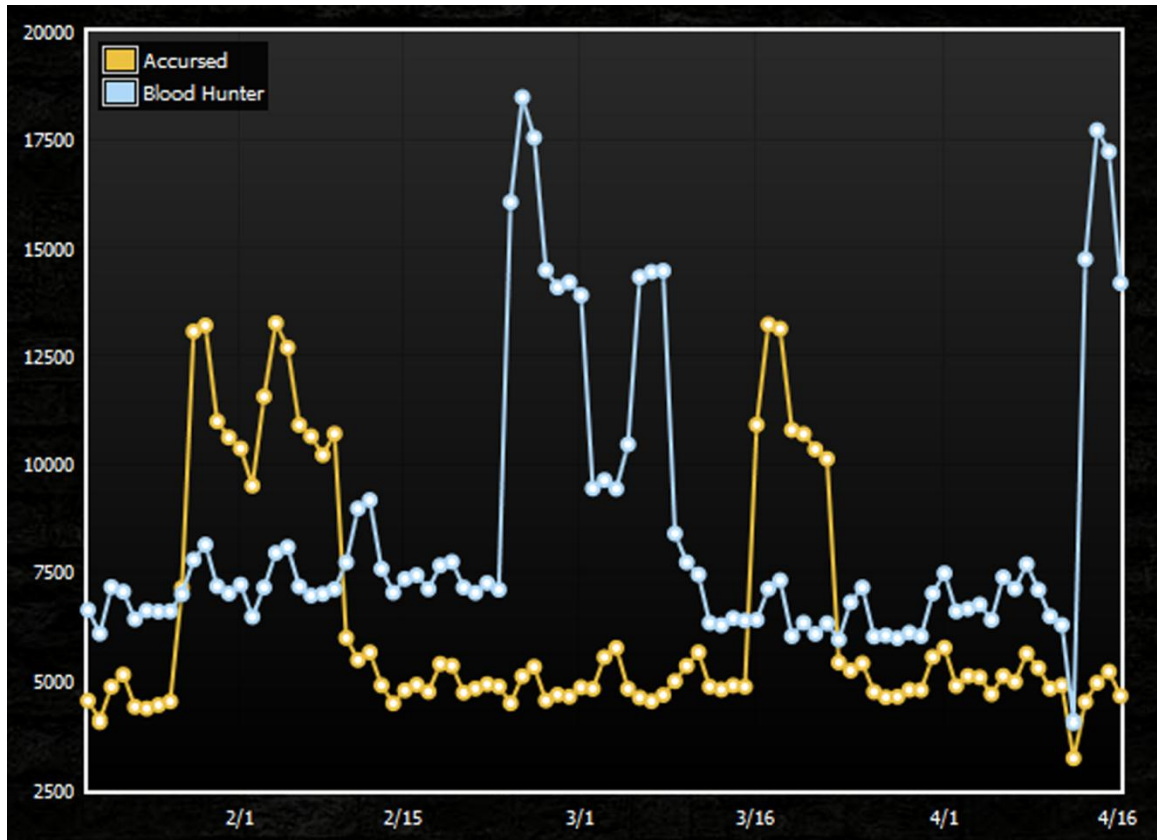


Figure 5.28: The graph shows how often two characters – the Accursed and Blood Hunter – were used by players in Heroes of Newerth over a three month period.

Global Statistic Examples

Simple examples of global statistics are often found on the homepages of a game's website. Websites for Bungie.net, EA sports, Infamous 2, and Killzone each have a few global stats declaring the size of their player community. They may say how many players are online, how many matches were played today or how many missions have been created. These values are more for advertising each game's community, more so than providing valuable data to analyze.

More involved accumulated and usage stats can be found in games like Battlefield 2, Driver SF, Halo Wars, Heroes of Newerth, League of Legends and Rockstar's Social Club. Here global variables are listed in regards to the completion of individual missions, the number of kills earned with a particular weapon or character, the favorite car or item used by the community or which map is played the most. While these global stats can be seen as advertising the size and strength of the player community, they at least show a nuanced rendition of how everyone within a game's community is approaching the game, even if the data is averaged overall. Some older systems built for Half Life 2 and Day of Defeat are exclusively global, no other type of stats are provided. Half Life 2's system also provides heat maps revealing where players die within the number of level that exist in the game. The game is single player only so the maps actually show, overall, where players have a difficult time in the game, based on where the players are dying.

Global War features are perhaps, along with community events, that add value to a player's gameplay beyond stats that represent the entire player community as a single mass. A global war is a meta-war on top of the individual "wars" being fought in each game's matches. These systems attempt to give value to each of those individual matches by having each of them contribute to the larger meta-war. When specific factions win matches in games like Brink or End War, those wins combine with other match results to determine which faction has won the most matches within a given area (usually represented by single maps in the game). When a faction wins more games in an area they are marked as 'controlling' that area and, therefore, controlling that portion of area as part of the larger war effort. Global stats become part of the game and can affect which maps players play (for instance if a player's faction has fewer wins in a specific area then players may try to play more matches in that area to even the win results) Although, war is not always the meta-game being played. In Assassin's Creed 2: Brotherhood players can purchase shops around the in-game city (i.e. Rome in the early 1500s) and based on how many players own a specific shop it may be more lucrative for players to purchase the shop versus another. If a shop is not owned by many other players, globally, then the shop may provide the player with more money. This shop buying system is similar to a global war system but uses commerce as the meta-game instead of war. When compare to

a global stat system found on a website for EA sports, which only shows how many players are online, some global stats can have significant effects on a player's gameplay experience.

Functional and Phenomenological Aspects of Global Statistics

Visible

Although global statistics collect data from every player in a game the amount of data ultimately made visible can be small. Global war systems for Brink or End War only release data related to which factions are winning more individual matches. Bungie.net, EA sports, Infamous 2, and Killzone only reveal a few high level stat totals like total number of players, total number of opponents killed or total number of missions created. A few systems only display global stats, such as Day of Defeat or Half-life 2, and data that would otherwise be defined as player data is averaged over the entire community of players. Regardless of the amount of data that is release, global stats are always filtered based on the sheer volume of players that contribute to the stats. No individual player's data is made visible and the stats instead represent either totals or averages of the player community as a whole.

Comparability

While some forms of global stats are merely total values accumulated over time, and therefore not necessarily comparing the change in values over time, other variations of global stats can be used for comparisons, albeit at a global level. Global war and community events can be setup in such a way to compare large groups of players to one another. Community events may be offered across servers or platforms, and players within each group may compete to see who can finish the event first (the WoW Ahn'Qiraj War Effort' is an example). Global war especially compares multiple player factions with one another because those systems are built around the factions trying to win more matches the others. Finally, usage statistics compare game content rather than players. Heroes of Newerth's system, for instance, can compare how often in-game

character and items are used over time while also providing comparisons related to other values such as how often players win when using a particular character.

Control

Global statistics are regularly controlled by a game's developer because they are the only ones capable of collecting data from every player in a game. A single player, or group of players, may be able to track some data within a game (for example, the Darkfall Political Map system is crowd-sourced data) but outside parties cannot track every match played in a game without being granted access (perhaps through an API) to the data by the developer. There is also the reality that a game will eventually lose most of its players and a lack of players turns "global statistics" into static values. Variation like accumulated statistics, usage statistic and global war are the most susceptible to dwindling player populations because these variations thrive on new data coming in to make their existence meaningful. For example, if there are only a few match results funneling into a global war system then the system loses the ability to act like a large meta-game being played. In fact, every global war variations found in the systems studied are either broken or have been removed, most likely due to dwindling player populations and the cost of system upkeep.

There are possibilities for outside systems to take advantage of, and control, some global stats. One example is the Noby Noby Boy global stat tracking system. A player created the system to web-scrape data (a way of taking data from a website by parsing the website's code) from the official Noby Noby Boy website that contains data related to how much length is added to 'Girl' each day (Noby Noby Boy is a game based on growing a worm like character, called 'Boy', and after each game the length of the Boy is added to a length of Girl, which represents the combined length of Boy from every other player's game). The Noby Noby Boy stat tracking system has scraped data from the website for the past three years and is able to show the progression of Girl's length over time. Although, the only way the system was built was by taking data from the developers website. If the developer of Noby Noby Boy ever removed the data from their

website the stat tracking system would cease to function. Developers usually have control over global stats, even when players find ways around collecting the data.

Exposure

Global statistics are typically too filtered to be seen as exposing private data related to individual players. None of the stats portrayed reflect directly on any individual player because global stats accumulate across all players. Global war features accumulate the results from player matches and community events accumulate resources and event frequencies from players. Although, exposing global stat data does make the player community more present. Exposing data about the community as a whole is a way for game developers to promote their game and show off the number of players that exist in the game. When an global accumulated stat on the Infamous 2's website exposes that a hundred thousand missions have been created by the community the stat is an advertisement to other players stating "our community is large and creating extra content for you". This helps players realize the size of a game's player community, which can sometimes go unnoticed when players only interact with a small portion of players over the period of the time with a game.

Interconnected

Even though global stats are regularly accumulated across all players, making the data less personal, this does not mean the stats lose their ability to connect players. Global stats connect players to the player community at large. Usage statistics, for example, can connect an individual player's play style with the whole player community. Players can see if the characters or items they choose during gameplay are used by a majority of other players, while also comparing themselves against the average stats of the community as a whole. Global war features also connect players to their game factions and they become part of something larger when their faction wins are added together. Community events too, make it feel like the entire community of players is working towards a common goal. Players do not connect to each other on a personal level when they use global stats but connect to the larger player community, to something larger than themselves.

History

Global statistic variations like accumulated stats or community events rarely provide a sense of history, although they each represent two different temporal representations. Accumulated stats represent a player community over time. Spore lists the total number of player created game assets in the tens of millions along with other global stat systems found in games like Infamous 2 and Little Big Planet. Those accumulated stat provide a way for a game's developer to acknowledge the size of their community but they do not divulge any changes over time. Community events also increase over time but those events eventually stop; they only provide an in-the-moment event that players can take part in. Accumulated stats can stay relevant after players have left a community, for example the Spore assets are still available to any future players, but when community events are over, they are scraped and the game moves on.

The other global statist variations, global war and usage stats, do have a sense of history (or at least can have a sense of history). Global war variations can maintain the results from past "wars" are maintained. End War, for example, was reset every week when the system was operational but players could revisit past weeks to see how each faction performed. The system allowed players to replay each week's war, visualizing how different nodes on the system map were taken over by different factions. Usage stats too can sometimes be temporally based. Heroes of Newerth provides graphs detailing usages stats for character selection, win rate with said characters and a way of comparing characters over the past few weeks. Even though the data found in both the End War and the Heroes of Newerth system has no personally important information for any player, the systems still provide a way for players to revisit past data related to each game's player community.

Global Statistics Summary

Global statistics turn the entire player population of a game into a single cohesive unit by combining data collected from every player. Sometimes this combined data is used as a means to market a game. Infamous 2, My Killzone, Spore all present global stats on their main home pages and are meant to show how large their player

communities are. In the case of games like Spore, which allow content generation, listing how much content players have created is a way of showing potential player the amount of extra content they receive if they choose to buy the game. Global stats can also connect players to larger events or meta-games. Community collection events ask every player from the community to pitch in resources or perform certain tasks in-game in order to unlock game content. Global war features collect the results from all the matches being played in a game and associate those results with a large war being waged between different in-game factions. For instance, Brink has two in-game factions and after every in-game match the faction that won would add their win to the Global war system. The faction with the most wins over time is marked as taking over different pieces of territory on a larger map. Global stats are both a way for developers to show the size of their player community but also act as a way players can contribute to meta-level events as part of the community.

Maps

While there are many types of visual representations used frequently within the play analytic systems surveyed (bar charts, graphs, etc.) maps are used often enough I decided to add it as a separate category. In fact, some play analytic systems surveyed only present data by using a map (i.e. Map WoW and Terraria Mapper), and therefore creating a play analytic category for maps was absolutely necessary. Maps are also the only form of visual representation that adequately describes the spatial context of data within a play analytic system. If a system plots the location of an item or the player within a game's environment maps are used. Considering many games render navigable virtual environments maps are a necessity to describe the location of in-game content and where events occurred.

23 of the 81 systems surveyed had some sort of map-based feature. The variations described next divide the type of map features according to the type of data presented on the maps. Sometimes maps are used to define where game content exists spatially within the game's environment; this includes marking where NPCs or collectible items reside. Other maps are used to plot player data and events. Player deaths, movement, or visited

locations are examples of player related data that play analytic maps display. Last, referring back to group statistic variation, ‘Global War’ systems almost always have some sort of global map they use to provide information about how each player faction fighting in the war are progressing.

Map Variations

The following map design variations appeared in the play analytic systems surveyed:

Game Data Maps

Maps used to display, spatially, where data referring to game assets and content exists within the game’s environment. Mapped game data is usually related to objects in the environment or locations. Objects like collectables, non-player characters, resource nodes and towns, along with locations like zones, appear on game data maps. These types of maps are often used by games with large open environments allowing players to explore the terrain non-linearly – Assassin’s Creed 2, Lord of the Rings Online and Terraria being some examples (REF see Figure 5.29). No player data is linked with the data being displayed on a game data map, unlike player data maps.

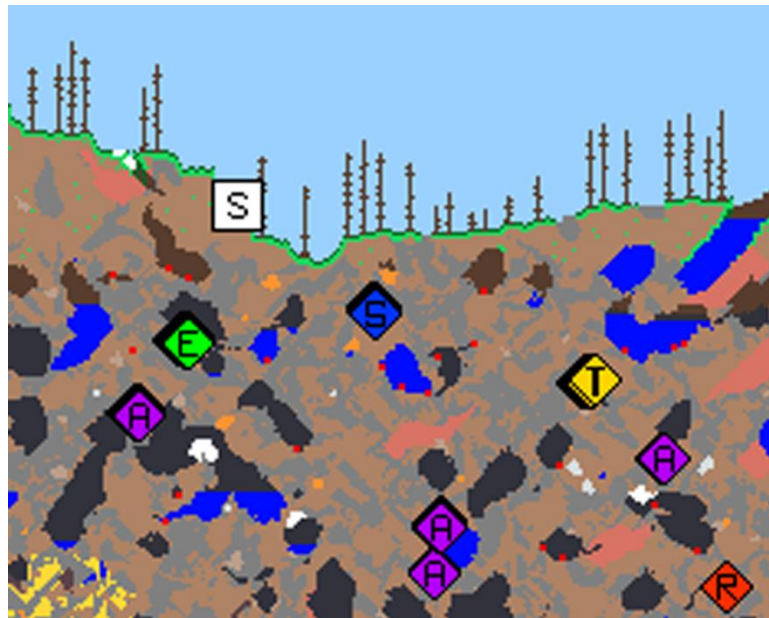


Figure 5.29: Some mapping systems create maps of a game’s environment based on a save file (Terraria Map pictured).

Player Data Maps

Maps displaying player related data such as player created events or locations owned by a player, or their group. For example, player data maps used in Rockstar's Social Club system (which provides data for the games L.A. Noire and Red Dead Redemption) display the in-game locations a player and their friends have visited (REF see Figure 5.30). Other player data maps are based on smaller, individual levels in a game. Gotham City Imposters uses maps of their multiplayer levels to show where the player kill and death events occur in each multiplayer match. Additionally, player data maps tend to be static or display only single, non-temporal, spatial points where player events occur. Only Bungie.net and SC2Gears was found to offer data related to actual player movement across time (e.g. a player's path during a capture the flag game of Halo 3 when the player had the opponent's flag in their possession) and only Halo Atlas offers real-time player positioning on a map.

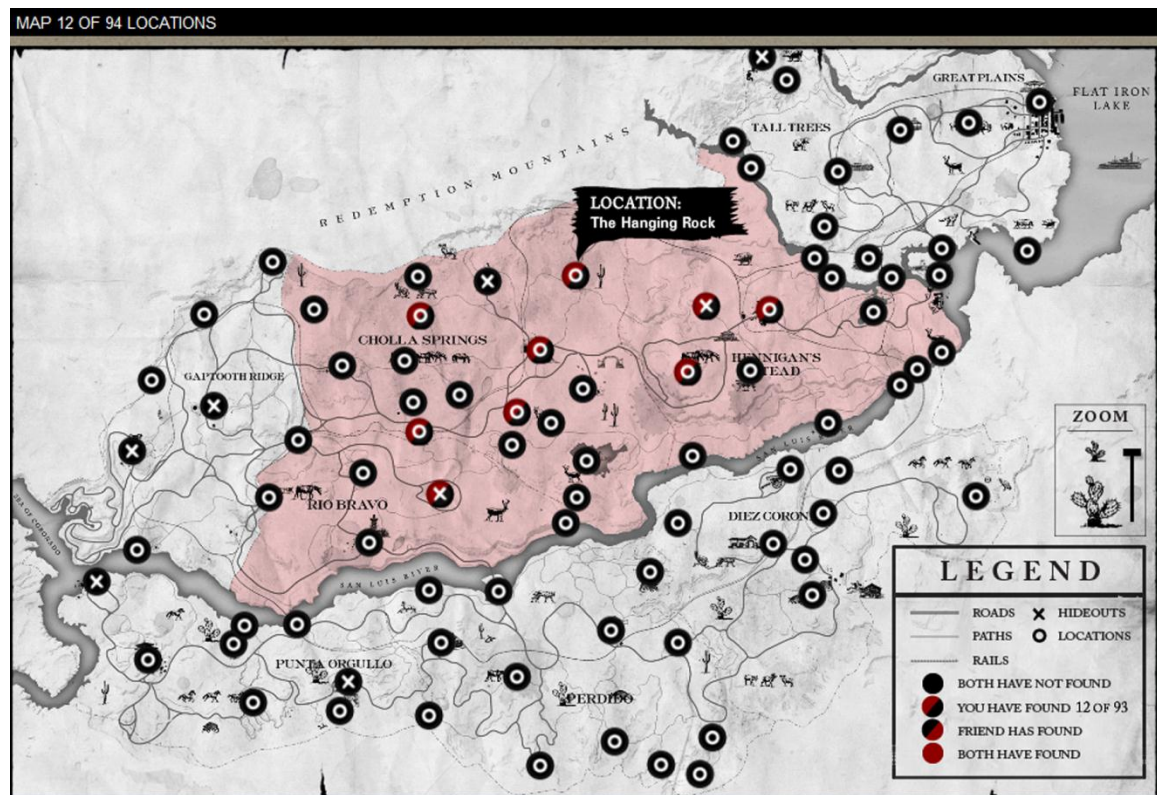


Figure 5.30: This map from the Red Dead Redemption play analytic system marks the locations where a player has visited and where that player's friends have visited.

Global Player Data Maps

Global player data maps are related to the global war variation of global statistics (see Global Statistics section in this chapter) and plot data representing an entire player population. The global war features found in games like Brink and End War uses a representation of a map meant to simulate a larger piece of terrain players are fighting over, but does not represent the maps found in the game (REF see Figure 5.31). This is unlike the maps used in Darkfall Political Maps and Planetside; both systems use representations of their in-game virtual environments to display data defining which player groups own the territories in the game. Similarly, Bungie.net uses maps of their levels to display heatmaps delineating data related to where players die and killed other players most often (these heatmaps can consist of a global dataset or data from individual players, in which case they are considered player data maps). All global maps display data related to player behavior on a massive scale, whether representing actual virtual environments or not.



Figure 5.31: Some game maps do not represent a game’s environment. Brink’s ‘Global War’ feature contains a map meant to represent the larger island players are fighting over in the game.

Map Examples

Play analytic systems based around maps tend to be built for games that have large virtual land masses that players explore. Map WoW, for example, is a player created interactive map system that presents the virtual environment from World of

Warcraft, using the Google Maps API. Users can pan and zoom into each of the large continents in the game while toggling on and off different markers on the map. The markers provide the locations of mass transit spots (e.g. flight paths), area names, resource nodes and treasure available throughout the environment. Although, Map WoW has not been updated since the latest expansion of World of Warcraft (entitled Cataclysm), an expansion to World of Warcraft that altered many areas of the game's environment. Since Map WoW is player created it is up to the original creators of the mapping system to keep the system up to date. This shows that even systems built outside of a developer's sphere of influence, and built by devoted players, can faultier over time, falling into a state of disrepair.

One way to keep a system from falling into disrepair is to make the users of the system provide the content for the system. This is how websites like WoWHead can keep content up to date and novel by crowd-sourcing their users. One play analytic system that builds game maps based on crowd-sourcing players is the Darkfall Political Maps system. The system itself keeps track of which guilds in the game Darkfall own certain towns or cities in the game. A political map – a map tracking where each Darkfall guild's influence resides on the game's landmass – is created each day to reflect the changes from the day before. One guild may take over a city one day and lose control over the city next week. Each event is tracked by the system and updated as they are verified. The events need to be verified because Darkfall Political Maps does not have a direct link to any data from Darkfall itself. Instead, it relies on the Darkfall player community, who also use the mapping website, to submit changes to the map. If a guild losses a city to another guild, a player submits the event to the political map system and other players must vote whether the event is true or false. If enough players vote that the event is true, the event is added to the historic log of city ownership events and the next day's map is updated. Unlike Map WoW, the Darkfall Political Map system survives based on the will of the Darkfall players using the system, instead of relying on a few players to maintain the whole system. If no one used the mapping system then the system would fail but then no one would be using the system anyways.

Finally, there are examples of maps that do not represent actual terrain or spatial environments from a game but use data collected from players to alter a map representing an abstract spatial environment representing a global war. I discussed the ‘Global War’ variations under the Global Statistics section in this chapter. Global War systems often make use of maps but they do not always use maps representing the actual terrain that exists in a game environment. Both Brink and End War, Brink being a first-person shooter and End War being a real-time strategy game, had a Global War system as part of their play analytic systems (both systems are no offline or non-functional) and did not use the maps used in their multiplayer matches. Each game’s system used an abstract map to symbolize a large piece of territory players are fighting over but the players never fight on this larger piece of territory while playing. The story behind Brink, for instance, pits two factions against each other as they are fighting for control over a floating, man-made island called the Ark. Each multiplayer match fought in Brink takes place on a map that represented a small piece of the overall island. Every time a faction won a multiplayer match the win was recorded for that individual piece of the island that the match was fought over. Throughout a set timeframe, say a week, as more match results were totaled for each faction, the faction with the most wins at any time was said to be the owner of that piece of the Ark. Over time one faction may control some pieces of the Ark but may be losing other pieces to the opposing faction. The map was there to represent these shifts in power and to give each individual multiplayer match more weight, because each match was affecting the ‘global war’. End War had a similar setup where three factions were fighting over real world pieces of land in Europe and North America. Over the course of a week different nodes across the two continents would change ownership based on which factions won the most battles within those nodes. Even though players only played on smaller maps, the nodes symbolized each faction attempting to take over continents as part of a larger war effort. Maps in play analytic systems, therefore, are not always directly correlated to a game’s environment and in some cases are used to provide an extra meta-feature, giving players a wider perspective on the effects of their gameplay.

Functional and Phenomenological Aspects of Maps

Visible

Maps, being a spatial form of visualization, make spatial data from games visible. Spatial data such as an area's name, the spot where a NPC spawns, an area where specific enemies appear, the location of collectable items, or a place where a player died. Any game-related data with a spatial location can potentially be placed on a map of the game's environment. There are also cases where data is projected onto an abstract map, one that players do not actually experience during gameplay but is meant to represent a larger area where the story behind the game is taking place (as in how Brink's and End War's Global War systems are setup). Spatial data is supposed to orient a player within the game's environment by provide information as to where to find game content or show where a player has been.

Comparability

Maps help visually compare where events in a game occurred or where to find particular assets in the game's environment. For instance, Red Dead Redemption uses a map to mark places a player has visited in the game and marks where a player's friends have visited, allowing for comparison. Maps of Halo multiplayer levels are used on Bungie.net to show where kill and death events occurred during multiplayer matches. Sometimes these kill/death maps are combined over multiple matches and a heatmap is used to mark the areas where kill/death events occurred the most. Darkfall Political Maps, Assassin's Creed 2 Brotherhood and Planetside all use maps to show which group of players own certain in-game territory found in each respective game's environment. Map WoW, Lord of the Rings Online and WoWhead each use maps to plot where specific NPCs and items exist in their respective game environments too. Maps are used in many play analytic systems to compare the spatial location of similar events together in order to find patterns of the event occurrences (e.g. the kill/death events in a Halo match) or to aid players in finding content in a game's environment (as a system like Map WoW provides World of Warcraft players).

Control

The location of the player and objects in a game's environment tend to be controlled by the game's engine. Rendering objects in game is handled by a game engine

and data regarding the location of in-game objects (players, items, world geometry, etc.) can be managed by the engine as well. Spatial coordinates of in-game objects, particularly the player, also change very rapidly throughout gameplay making it harder, and more costly, to track the spatial data over time. Developers who use maps in their play analytic system routinely record a limited number of spatial locations of certain events or objects because the benefit for storing every change in spatial position is limited. For example, the map used in Red Dead Redemption tracks where the player has visited, i.e. what key areas they have been to, and does not have to worry about tracking how the player traveled in-between those locations. Players may have spent hours wondering in-between those key areas and, while it is arguable whether that data is valuable to the player or the developer, would be costly to store over time. Additionally, since a player's position data is rarely recorded over time it is hard for outside parties to get ahold of any position data from players. This often makes developers the ones who control how often and to what extent position data from players or objects is released.

For outside parties to gain access and control over spatial data they resort to other means besides relying on developer to provide them with spatial data. One way outside parties, particularly players, can gain access and control over spatial data from a game is by access replay or save game files. These files typically store data pertaining to spatial coordinates relevant to a game's engine. If players can reverse engineer a replay or save game file they gain access to the spatial data. Minecraft X-ray, SC2Gears and Terraria Maps are all examples of systems where players reverse engineered game files to unlock the spatial data inside and used the data to build mapping programs other players can use. Another way outside parties can gain access to spatial data is through build game MODs (game modifications). WoWhead, for instance, asks players to use a tracking add-on, that is a type of World of Warcraft user interface modification, during gameplay in order to collect data from World of Warcraft. Spatial coordinates of items and enemies are recorded and sent back to WoWhead to be aggregates with other player's data. The data is then used to produce maps for those in-game assets, describing where in the game environment players those assets are located. Finally, there is always the option to collect spatial data manually as some player created play analytic systems do. Darkfall Political

Maps uses a crowd-sourcing model where Darkfall players submit changes indicating guild ownership over the many towns and cities in the game. As towns change hands between different in-game guilds, players vote on which changes are true. If a submitted change earns enough positive votes the political map is updated the next day. Spatial data may be hard to gather without some intervention by the player (whether reverse engineering a file, building a MOD or collecting data manually) but once that intervention occurs players often build highly useful play analytic systems.

Exposure

All game maps bracket data within a game's environment, connecting spatial position to content or events, but the level of data exposed is different from game to game. Systems like Map WoW or WoWhead expose the locations for many of the items, NPCs, quests, etc. found in World of Warcraft. These maps work well with content databases, such as WoWhead, where location data aggregated from many players can be compiled into one map. Exposing the locations of game content also takes away the exploration aspect of playing the game. Players learn where the best locations for finding certain items are and do not go through the process of deducing that information themselves (some players think providing such detailed location data for a game is cheating (Consolvo, 2007)). The maps used for Red Dead Redemption and L.A. Noire expose where a player has been and where their friends have been. Halo Atlas even exposes a player's position in a Halo: Reach multiplayer match in near real-time using a mobile device. These maps make it easy to monitor a player's progression through the game's environment but also makes it hard for players to hide where they have been, in case they wanted to keep that data private. Autolog and RiderNet use maps as the selection screen(s) for choosing races. Each race is attached to a friend leaderboard, which appears on the map interface, and allow players to monitor their friends' high scores. Each of these examples show how maps are used to expose different types of data in a variety of ways. However, the main divergence between maps and other play analytic categories, such as leaderboards or statistics, is that maps can expose data referring to a "physical" location within a game's environment. Most maps do not provide abstract representations of data like total number of kills or high scores (although global player

data maps are close), and instead expose specific locations where in-game events occurred or content can be found. Exposing data related to in-game locations becomes much more personal because players actually exist within the environment the data indicates. Players can visit the locations plotted on a map, which provides a different experience when compared to viewing abstract statistics.

Interconnected

Maps connect players spatially to the events that occur in a game's environment and to the environment itself. For games like Terraria that do not have an in-game map the Terraria Map Viewer allows players to analyze their game environment and plan ahead. Special items and resources are marked on the map giving players indication of where to dig to next and allows them to work towards collecting resources they are after in the game. Maps can also link friends together and help them monitor each other's progress. Red Dead Redemption and L.A. Noire both use maps that reveal a players progress through the game environment and the progress of their friends. It makes it easier to know how far their friends have gone in the game and may encourage players to ask their friends about specific areas in the game. Players can connect to other players competitively using maps as well. Bungie.net and COD: Elite provide maps of multiplayer matches and plot the locations of each player's kill/death events. Players can reflect on these events to determine if their gameplay strategies are working based on where they are getting kills and being killed by other players. Maps allow players to analyze a game's environment and their own actions, and the actions of other players, in that environment.

History

Maps tend to be biased towards offering the most up-to-date data possible without acknowledging the history of the data collected. WoWhead, for example, provides maps detailing the locations where players can find most of the NPCs that exist in World of Warcraft. The data available has been collected over many years but, because the NPC locations rarely change over time, the data is presented as if it was added all at once. Other maps such as the heatmaps provide by Bungie.net, Half-life 2 and Just Cause 2

(each marking the kill/death locations of players) are also examples of ahistoric maps because they combine data from many players to form aggregated views of where players are killed within each game's respective environment.

When maps are built to show data across time they plot varying amounts of data. Gotham City Imposters and COD: Elite use maps to plot the locations of events over each multiplayer match but only a few types of events are shown (i.e. kill/death and objective completed events). End War used an animated world map for its global war feature (see the Global Stats section above) and would track when player factions took over certain nodes as part of the global war meta-game. Each week was a new "war" and the map kept an historic record of each week's results. Although, global war features do not provide data regarding individual games and aggregate data from all the players in the game. Therefore, the data presented on the End War map is only relevant to the community of players taking part in the war during the week it was recorded. Finally, SC2Gears perhaps has the most historically focused map feature with the ability to view a Starcraft 2 match across time. Players can replay a match over time and witness where: buildings are placed, units are moved to and battles take place. SC2Gears is able to almost identically reconstruct Starcraft 2 matches by using the replay files that are created after each match concludes. Many games do not create or have the ability to save replays, making it difficult to build something like SC2Gears for other games. It is possible, however, for maps to provide a historic record of the events that occur in a game; there just needs to be a system in place to record a game over time.

Maps Summary

Maps in play analytic systems are both used to cart the topological aspects of a game environment and mark the location of content or events that occur within the environment. For systems like Minecraft X-ray and Terraria Maps, players can examine the entire topology of their generated game environment (both Minecraft and Terraria generate a new world for each player) that is stored in their saved games. This includes plotting where resources exist, dungeons are located and where to find treasure. Additionally, Play analytic maps regularly contain features that are not available to

players during gameplay. Bungie.net and COD: Elite provide players with heatmaps of their kill/death events during multiplayer matches. Darkfall political maps keep records of which player guilds control certain cities in Darkfall, a feature not provided in the game, and WoWhead marks the location of every NPC in World of Warcraft. Other times maps provide data that is available to players during gameplay but are hard to access. Map WoW, for example, marks where all the mass transit locations in World of Warcraft are located, which can only be accessed in the game when a player is at one of the transit locations. Halo Atlas mimics the type of map system that most games offer during gameplay, real-time player positioning including the location of relevant data such as items or other players, but is meant to be used on mobile devices detached from the game's run-time state.

Players experience a game's environment using without actually running the game when they use play analytic maps. They provide an almost tactile experience as players can pan, zoom and explore a game environment from a bird's eye view. Even when play analytic maps are presented as part of the game's run-time state these maps are removed from the actual gameplay experience but provide a chance for players to reflect on their gameplay (often through monitoring the friend leaderboards available in the Autolog and RiderNet systems). It is experiencing the game without the burden of gameplay. Other play analytic categories like player statistics or leaderboards do not recreate a game environment in order to provide data, they are merely scores. Maps forces an intrinsic link between a game environment and the data being spatially represented. This makes the data both more exposed, events are placed spatially (sometimes temporally as well) in addition to being tracked, but better capable of interconnecting players with the game environment and the events that take place there.

Content Generation

Creating content can be an essential component to how a game functions and is experienced. Games like Little Big Planet and Mod Nation Racer are built upon players creating game content. Both games provide players with robust creation tools that allow players to model and adjust their own levels and race tracks. These games thrive on the

willingness of players to create game content and players are more than happy to create content when they find the activity fun. The issue that arises when players start creating game content is how the content is distributed to a game's player population at large. A game can provide creation tools, or players may create game content using their own tools, but in order for the content to have value to the player community there must be a way to transmit the content to other players. This is why play analytic systems, in addition to collecting and presenting statistics, often provide a method for distributing player generated content to a game's community at large when creating game content is a big part of the game experience.

21 of the 81 systems covered in the study had some form of content generator and sharing feature. These systems act as both a show room floor, displaying the created content, and a filtering mechanism, allowing players to rate and describe the available content. Rating content, i.e. acknowledging the quality of a piece of content by scoring the content along a fixed scale, in particular is ubiquitous amongst the systems, even though it is not necessarily a ubiquitous feature in all user generated content systems (like wikis, forums or blogs). Rating abilities allow the play analytic systems to filter content based on quality, or at least filter content by what a majority of the player population believes to be of high quality. As the variations within content generation are discussed below, it should be noted that every system applies a rating in some way.

The variations presented below are based on where content is generated by players. Games with built in creation tools allow players to create content while playing the game and can upload their creation to the play analytic system (e.g. Little Big Planet and Mod Nation Racer have in-game creation tools). Some play analytic systems have online creation tools that allow players outside of the game to create content, share it and download it into their game. Skate 3 has a creation tool for both creating avatar images for players and texture images they can place on their in-game skate boards or skate parks. Other times content is created that has no connection to the game experience, meaning the created content is never used in the game. The Sims Exchange allows players to create movies or slideshow stories about the sim families they make in the

game The Sims 3. The stories use video or screenshots from the game but the stories are never incorporated back into the game, they only reside in the play analytic system. Finally, content creation and statistics often merge within a play analytic system. Levels built in Little Big Planet have their own leaderboards for example, and I have included a variation on content generation to point out when the link between statistics and creation merge within a play analytic system.

Content Generation Variations

The following content generation variations appeared in the play analytic systems surveyed:

Game Content Sharing

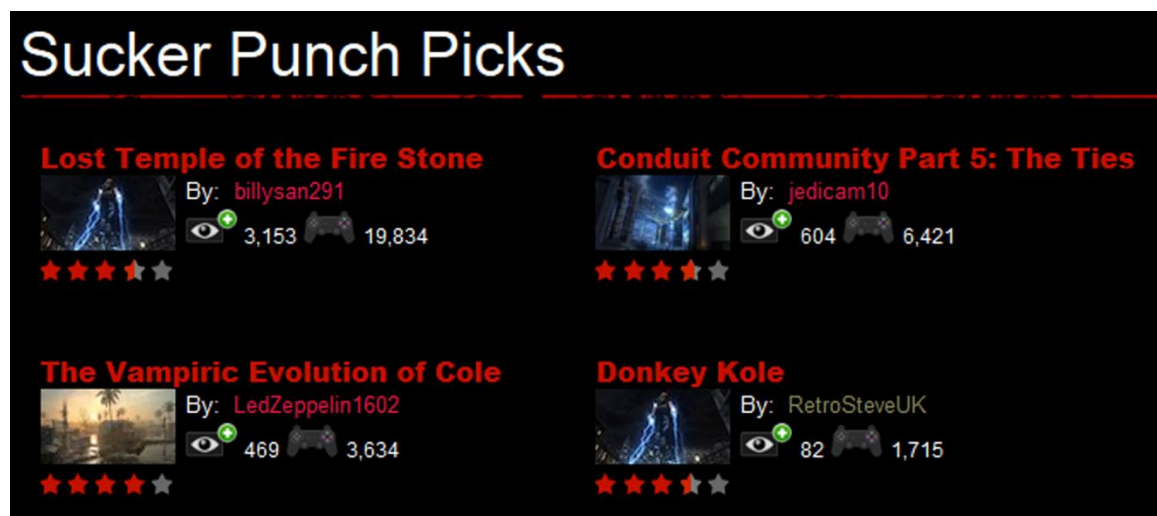


Figure 5.32: Infamous 2 players can share levels they create in the game online and other players are given the chance to rate the levels for quality.

Tools for creating content are sometimes provided alongside a game as part of the game experience and shared amongst the player community. These tools may allow players to take screenshots, capture video, create avatars, model 3D characters, build levels and design textures. While these tools are setup alongside the game experience, and do not necessarily have to be taken out of the game itself, the content produced using these tools often make their way to the game's play analytic system. There, the content is organized and made searchable for other players using the system. For example, Infamous 2 players can create missions in the game and the missions show up on the

Infamous 2 play analytic system where players rate and comment on the missions (REF see Figure 5.32). Game content sharing, thus, refers to content created in a game, and is available in the game, but is moved from a game to a play analytic system.

Game Content Creation

Content produced in the play analytic system is shared amongst a player community and is made available in the game. Instead of players using tools in a game to create content, some play analytic systems offer their own tools for creating game content. The play analytic systems for Army of Two: 40 Day and Skate 3 both have texture creation tools (REF see Figure 5.33) players can use to make symbols for their in-game characters or other objects in the game (for example skateboards for Skate 3). Content created in a play analytic system work in the same way as content created in-game. The content is made searchable for the community of players who in turn rate and comment on the content. Content created in a play analytic system, however, goes in the opposite direction compared to game content sharing; content is made outside of a game and is then made available in the game.

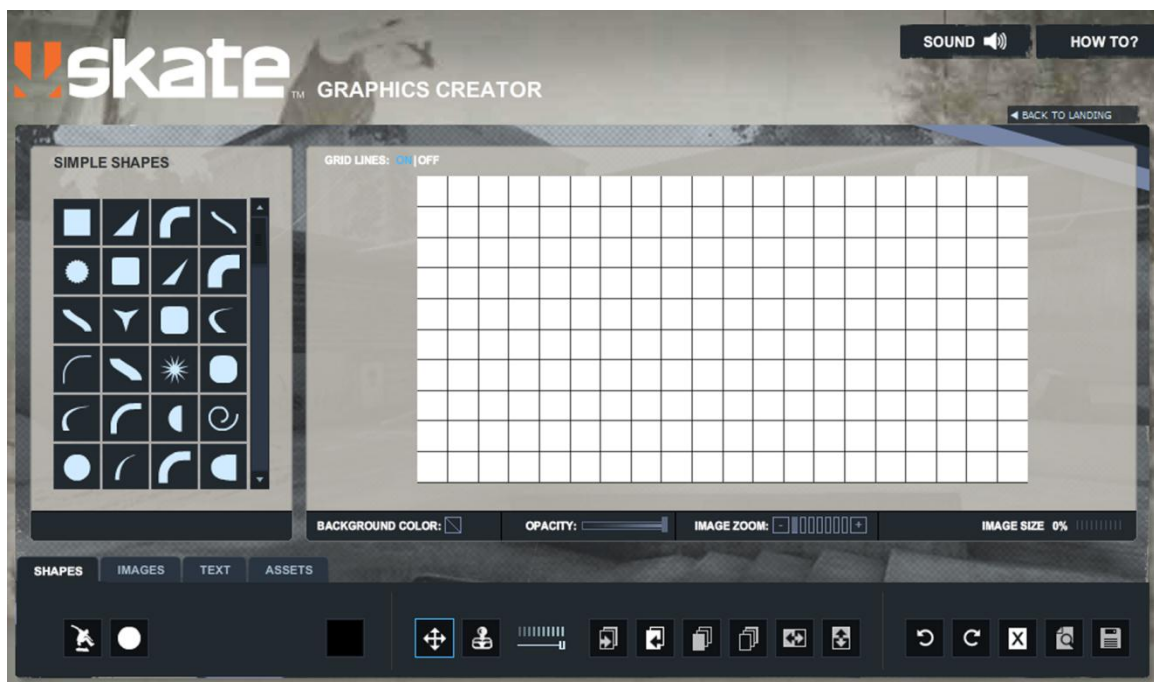


Figure 5.33: Some play analytic systems have content generators built into the system. Skate 3's 'Graphics Creator' afford players the ability to create their own icons and graphics, which can be used both in the play analytic system and in the game.

Non-game Content Creation and Sharing

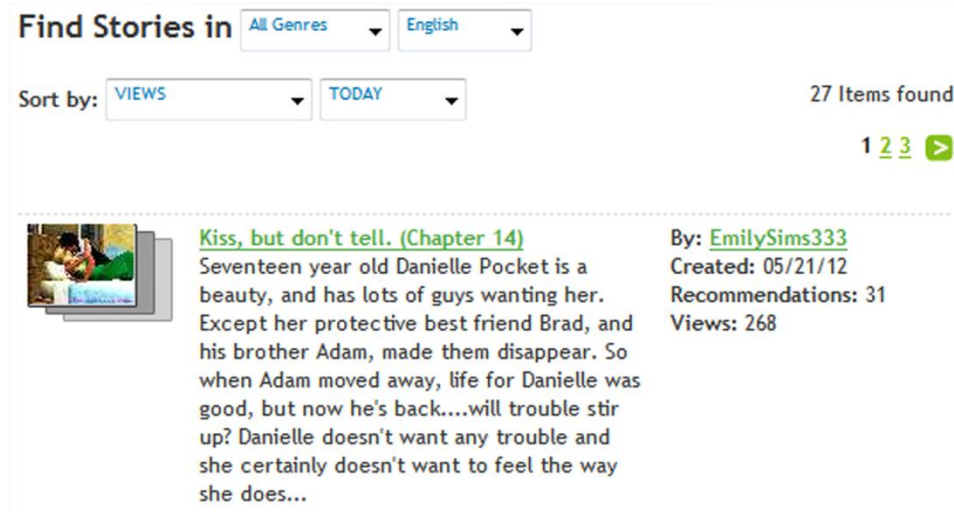


Figure 5.34: Some player generated content is not meant to be used in a game environment. The Sims Exchange allows players to create stories using screenshots from the game but the stories are only available on the Sims Exchange and not in the Sims 3.

‘Non-game Content Creation and Sharing’ refers to content produced that is related to a game but is not produced in-game or used by the game after the content is created. There are a few examples of play analytic systems offering tools to players for creating content that is related to a game but is not actually used within the game. The system built for Bad Company 2 offers players a strategy creator, for example. The strategy creation tool provides players with top-down images of the game’s many playable maps. These maps represent the game environments where in-game matches take place. On any map image, a player can lay down lines, icons and text to describe a strategy for completing the objectives on the map. After a player marks up a map and finalizes the strategy they can share the strategy with friends. At no time, however, is the strategy used in the game, it is strictly content that is produced, and shared, within the play analytic system. Similarly, the Sims Exchange allows players to create their own story slideshows (REF see Figure 5.34). Players take screenshots from the game and add captions to the images. These captioned screenshots are strung together and players use the tool to write stories about their Sims characters. Even though the screenshots come from the game, the stories created are unique pieces of content that are only created on the Sims Exchange and never used in the game. Play analytic systems can therefore

include the ability to create content related to games as a way for players to participate further in the community and help them enjoy their gameplay in other ways besides playing a game.

Gameplay Analysis and Sharing

Extra statistics are gathered from content, produced either inside or outside of a game, related to gameplay. Some types of player created content are related to gameplay: game levels, saved match replays or full games (as is the case on the Flash-based gaming portals like New Grounds). Content of this nature can work differently compared to other content like videos or textures. A created level can be played by other players and that means statistics can be gathered from players playing that level. For example, Little Big Planet allows players to create levels for other players. A level can be setup to track how long it takes a player to finish the level or track how many points a player earns while making their way through the level. Each player created level gets a leaderboard and by using either the temporal or score data, players are ranked on each level's specific leaderboard. Created content that is playable therefore becomes more than an item a player uses or a video a player watches; playable created content, such as a level, becomes another object for analysis. Replays recorded from games are another example of content that is analyzed, for instance SC2Gears and SC2Replays.com are systems built to analyze the gameplay found in Starcraft 2 replays. In the case of SC2Gears, the replay analysis system is considered to be a type of match stats but sometimes new content is created based on gameplay analysis, such as SC2 Replays' "player dna" feature that visually depicts an abstract version of a Starcraft 2 player's build order (REF see Figure 5.35). The "player dna" is based off of multiple match replays collected from a player, and could be considered a player stat, but the "dna" image's compact size means it can be shared with other players like a piece of created content (for example, the dna images acts as a texture or avatar image a player makes to represent them within a play analytic system). Finally, the Flash game portals (Armor Games, Kongregate, NewGrounds) treat individual games as separate pieces of content. Each game offers players their own achievements and leaderboards; in this case a whole game created by a user of a Flash portal acts like a piece of generated content where a player's gameplay is analyzed.

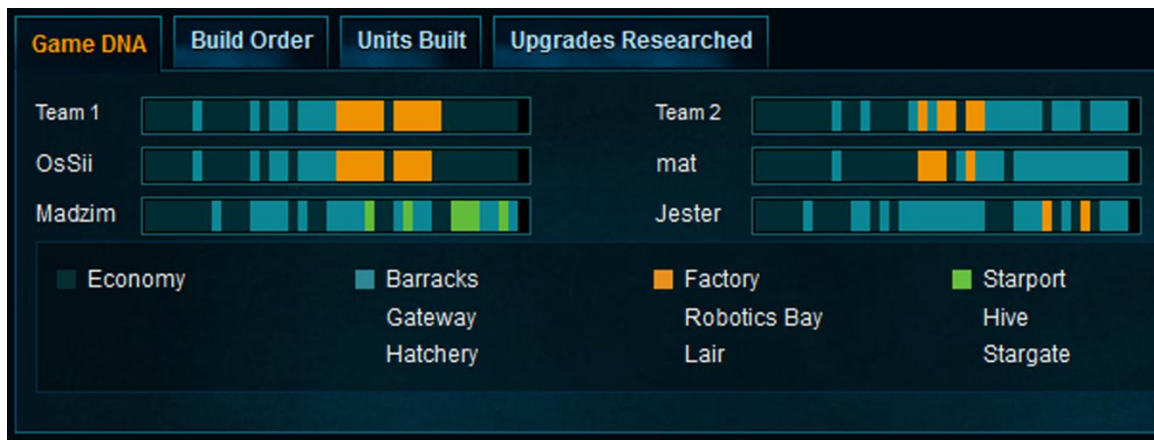


Figure 5.35: Some generated content is eligible for further analysis because the content is playable or consists of player data. SC2Replay.com analyzes Starcraft 2 replays and create ‘Game DNA’ images that represent the build order of the participating players in the replay.

Content Generation Examples

Content generation features allow players to create their own game content and share it with a player community. Both creating and sharing content can happen in two places: inside a game as part of the game’s system or outside of a game as part of a play analytic system. Spore, for example, contains a tool called the creature creator available to players while they are playing Spore. The creature creator is a limited 3D modeling and manipulation tool players use to build creatures, buildings and spaceships that appear in their game. Players are given tons of individual parts and pieces that can be attached together to form various models. Once a player has finalized a model – let us say they produced a creature for instance – it is uploaded to a player’s personal Spore account as one of their creations; acting as content created inside a game that is moved to a play analytic system. Outside of the game, if the player visits their Spore player dossier they can find a list of every model they have uploaded. Players can then share their created models with the Spore community by marking the model as public, giving it a name, adding a description and tagging the model with keywords. Adding the extra meta-data, like keyword tags, help other players search for models using the play analytic system online. When other players find an interesting model they may leave a comment, rate the model for its quality or add the model to their own game. Spore allows players to

download and add models created by other players to their game, encouraging players to share content with each other.

Other games follow a similar format compared to Spore's content generator feature, where players create game content in the game itself and then share them with the community online using a play analytic system. Little Big Planet, Infamous 2, Mod Nation Racer, and Bungie.net are all examples of system where content is generated in the game but available to an outside play analytic system. Little Big Planet allows players to create levels using an in-game level editor. Infamous 2 too allows players to create missions within the confines of the open world area that exists in the game's environment. Mod Nation Racer provides players with a race car modifier so players can create their own cars to share. Finally, Bungie.net has both screenshot and video capture functionality players can use to capture images from their game and upload them online. While it should be noted that most of these system allow players to search and share generated content through the game itself (meaning while a player has the game loaded) each of these systems also presents the generated content on each of their respective play analytic systems.

Sometimes content generation tools are provided to players as part of a play analytic online system instead of the tools being offered inside the game. Play analytic systems built for the games Skate 3 and Army of Two each have tools for creating icons and 2D images. The tools give players an assortment of shapes and symbols to produce the icons and images, much in the same way Spore provides players with pieces to produce 3D models. Players use the images they produce in Skate 3's and Army of Two's creation tool to act as their avatars, images that represent the player's identity within the play analytic system, and can be used inside the game, sometimes as labels or images that can be placed within the game's environment. Skate 3 allows players to place the icons on skate ramps or tracks they produce in the game, for instance. But no matter where the content is created, either inside or outside of a game, content generators always afford players the ability to share the content they generate and that content becomes part of a

player's overall experience with a game (either inside the game or outside the game in a play analytic system).

Functional and Phenomenological Aspects of Content Generation

Visible

Content generators are built to share game content. In other words, they make game content created by players visible. Player created game content is similar to statistic data because both forms of player data, statistics and created content, are shared using play analytic systems. One difference between statistics and created content, however, is the ability for other players to rate created content. If an entire community of players are creating new content each day, a play analytic system must have some way of organizing the created content to make searching for interesting content easier. This is done a number of ways. One, a numerical system is setup to rate the quality of a piece of content. Infamous 2 for example allows players to rate player created missions on a one to five star scale. Two, tags are attached to pieces of content and are used to describe the content. Little Big Planet's system allows players to tag levels created by the community. These tags can be descriptive of the gameplay (e.g. short, hard, single-player) or can describe the aesthetics of the level (artistic, ugly, scary). Three, comments can be left regarding pieces of content. Every creature created in Spore has a comment section related to it and players within the community use it to acknowledge the quality of the work or to ask the creator questions. Systems can use one of these methods to organize content or all three. While player created content is made visible in a content generation system, it is the organization methods, the meta-data attached to the content, that allows the content to be found, and judged.

Comparability

Comparing player generated content is typically left to the rating system used to organize the content. A typically rating system offers a sliding scale for players to mark their fondness for the piece of content being rated. A scale usually asks players to rate a piece of content along some range between one to five or one to ten. Players rate a piece

of content higher when they feel the content is of higher quality or they find the content personally meaningful. For example, one creature created in the game Spore was made to look like a human being. The game's creature creator is actually built in such a way that building a human is not a straight forward process but the player who created the model found a way. As a result many other player rated the model very high as it made it possible to add humans to their game. The human model may have been less sophisticated than other models but its uniqueness earned it a higher ranking. This also means that ratings can be subjective but the problems surrounding rating systems, as well as reputation systems (i.e. rating individuals), are well documented (Farmer and Glass, 2010). Regardless of their problems, ratings are the most frequent way content is compared within a content generation system.

Control

Game developers usually have control over the content generation system that manages and organizes the content being created by players. Developers decide how players can search for content and what meta-data is available for describing the content. The one thing developers do not have control over is what players actually create. However, lack of developer control over the content created is the point of these systems. Game developers often add content generators in order to entice players to create content for the game. New content keeps a game fresh and also evokes the creative drive of players. Additionally, if the generated content is mobile, meaning players can download and access the data outside of the regular play analytic system, they can build their own content management system. SC2Gears and SC2Replays.com are example of systems built by outside parties to analyze the replay files created from Starcraft 2. Spore Skeleton is an art project that makes use of the 3D skeleton data from player created Spore creatures, another example of an outside party building off of player generated content. Developers may control the creation tools in these systems but players control the creativity.

Exposure

Content generators are built by developers to give players a creative space to play and create content. This can help players stay engaged with a game because new content is always being added. But the only way the system works is if new content is continuously exposed in order to achieve the goal of engaging players. Although, exposing player created content is different compared to exposing data within other play analytic categories. When player statistics are exposed, for example, the player typically has no choice. If a player uses the Battlelog system for Battlefield 3 their data covering their recent matches and win/loss record are always put on display. Exposing content players create is different because a player has control over whether they release the content they create or not. Most of the content generators have a method of marking content as either private or public. My Steelport, a system built for Saint's Row 3, has a character avatar creator and players can share their created avatar with the player community. Players can also mark their created characters as private and no other player is given access to the avatar. Content generators, therefore, often have better privacy controls than most forms of play analytic categories. This may be due to the fact that statistics are not seen as a major privacy issue compared to releasing created content (even though a developer owns the content as much as they own the statistic data according to a game's EULA – end user licensing agreement). Or, perhaps developers do not want everything players create to be immediately public because players may wish to experiment first before releasing their final product. By having a public/private line players must decide if the content is ready for release, and hopefully this keeps the amount of sub-par content exposed to the community to a minimum. Whatever the reason, player content is the data being exposed within a content generation system and the players routinely have control over when their created content is exposed.

Interconnected

Content generators not only allow players to create game content but allow players to share their created content too. Players who are drawn to creating game content can find other players who enjoy making content too within content generators. The systems built for Bungie.net, Little Big Planet, Spore, etc. all have means for players to communicate and follow each other in the system. Most pieces of content have comment

sections for players to discuss how the content was made or comment on the quality of the work. 'Follow' features are often implemented in content generation system to make it easier for players to monitor other players. These types of systems provide both a way for players to create data but to also monitor other players who are excellent content creators.

History

Keeping track of new content is often a priority of a content generation system. Every piece of content created is given a timestamp and makes it easier for players to search for new content being added to a game at any given time. The newest content is typically highlighted within a system above all other content, for example new levels created in Little Big Planet are promoted on the main webpage when users browse to the game's play analytic system, 'LBP.me'. Promoting the newest content allows players who are currently creating content for the game to become more visible to the community but at the expense of past content. However, new content does not always trump high quality content. Most content generators use a rating system, or voting system, to define the highest quality player created content. Instead of using time as a filter for searching, players can also search for content based on the ratings the content has earned over time. Having highly rated content is important to the developers of the game; any content rated high means more players are likely to interact with the content and continued to be engaged by the game. Although, this supports a bias within the system because if quality content is rated higher it continues to be visible to the community and has a higher potential to be rated even higher. The highest rated content continues to be visible while other content is left hidden until a players specifically searches for it. Therefore, rating systems allow past content to become visible over time, where the highest quality content has the best chance to stay visible, but if a content generator promotes new content that content is given a chance to earn a high rating by being made visible alongside higher rated content.

Content Generation Summary

Content generation is about creating content and sharing the content with a game community. Sometimes players create content in game and the content is moved to a play analytic system where it is organized and made searchable. Other times, play analytic systems offer content generators as part of their feature set and players create content that is either imported back into the game or the content may only exist within the play analytic system. Players regularly have control over what they create and what they expose to the player community at large, even though many of the content generators are controlled by developers. Content generators also give players a chance to connect with other players who enjoy creating content and give all players the chance to rate and comment on content they find meaningful.

Content Database

While some games rely on players to create content as part of the gameplay experience, other games have so much content a play analytic system needs to be built to inform players about what content is available. This is particularly true when you look at MMO's like World of Warcraft and Lord of the Rings Online. Each game has hundreds of thousands of piece of content related to gameplay. There is such a variety of in-game items, NPCs, questions, locations, and enemies that these games run into the same problem as the content generators do; the system needs to organize and present the content data to players in a coherent way. Therefore, when a play analytic system is used to organize and disseminate content created by the game's developer these systems are known as content databases.

Only [ref 8 of the 81] systems surveyed had some form of content database but, just like how game leaderboards are more prevalent than what are found in the systems covered, there are other content databases I did not include. Most content databases follow a wiki style system where users can add data over time as they experience the data in a game. Wikia.com promotes itself as a community of game wikis, where virtually every game released is given its own wiki and players can add to any game's content section. However, Wikia is not covered in the play analytic systems surveyed because the

system as a whole is too large to cover. Other systems like WoWhead and Giant Bomb are covered in the surveyed systems and work in a similar way compared to Wikia by offering players the chance to add information about a game's content. Plus WoWhead and Giant Bomb also have other play analytic categories, like player statistics, which make them prime candidates for this study verses the more common examples of game content databases such as Wikia. Thus, while the number of systems within the content database category is low within the population of surveyed play analytic systems, most games have some form of content database created for them.

The variations of a content database are split into two types. The first type, as it is explained below, is the static version of a content database. Content added and disseminated to players is typically monitored by a game's developer and updated in chunks over time as new content is added to the game. This forms a static database because all of the available content is known ahead of time and is added all at once. The other type of content database is the dynamic database, which is often run by users. A dynamic database continuously adds data over time because of many factors: a) the game's developer does not provide data about the game's content and players must catalog the content themselves, b) content may continuously stream in and need to be regularly updated and c) players may wish to catalog content a game developers does not wish players to catalog (such as how to finish a quest quickly) and this data takes time to gather from a game. The difference between the two variations is really how the content being cataloged changes over time and how the owners of the database relate to the content (whether they are a game developer or player).

Content Database Variations

The following content database variations appeared in the play analytic systems surveyed:

Static Databases

Covers data regarding items, areas, story points, objects and other aspects of a game but does not change over time very often. Static content databases are usually

found in play analytic systems built by a game's developer. The content provided is meant to augment the gameplay experience and the statistics being gathered from players. For example, in play analytic systems built for first-person shooter games (Battlelog, Bungie.net and COD:Elite) content data related to guns, equipment and levels are provided to players. This data is presented alongside other statistic data referring to how many kills a player earned while using a weapon or how long it will take for a player to unlock a piece of equipment. The content data does not change over time, however. A description of a gun in Battlelog is always true and remains the same, unless a new game expansion is released. The same situation occurs for massively multiplayer games like Aion and World of Warcraft (REF see Figure 5.36). Data related to in-game items, questions and npcs are displayed as part of player's load-outs (see Player Statistics section in this chapter). Generally, static content databases are meant to provide players with standard information regarding the content found in the game and are meant to replace the need to provide game manuals covering all of the content available.



Figure 5.36: Static content databases, like the World of Warcraft Armory pictured above, are used as reference manuals for players to search of game content (e.g. items, quests, NPCs, etc.).

Simple Dagger



Simple Dagger

One-Hand Dagger

3 - 6 Damage Speed 1.80

(2.5 damage per second)

Durability 20 / 20

Item Level 6

Sell Price: 38 

Fishing Locations

This item can be fished in **Elwynn Forest** (27) and **Durotar** (18).



Figure 5.37: Dynamic content databases provide a similar service as static databases but often add additional data or player knowledge to the content data the database provides. For example, WoWehead displays map coordinates of items and NPCs based off of data collected by players who interacted with those items and NPCs in the game World of Warcraft.

Dynamic databases covers data regarding items, areas, story points, objects and other aspects of a game and changes over time based on player input. Unlike static databases, dynamic databases change over time either due to a consistent in-flux of new content. This may be a result of players having to discover content data not described within another static content database. Databases dealing with achievements are an example of systems constantly dealing with new content. Most games built for specific gaming platforms add platform specific achievements. The content databases covering achievements (achieve 360 Points and Giant Bomb) must continually add the achievements of new games into their system. In other cases, a database is dynamic because players add data they gather from a game that would otherwise be only available to players who experienced the data for themselves. For instance, a static database may describe an item, along with the item's stats, but does not describe where the item can be found. As players find the items in the game they can add the location data of those items to a dynamic database. WoWhead is an example of a dynamic database where World of Warcraft players continuously add data related to the game's content that is not shown on other WoW static databases like the WoW Armory (REF see Figure 5.37). A game's developer may not encourage these types of dynamic databases to exist because they want players to experience their game instead of searching for the fastest way to accomplish a task. However, it can be argued that these databases are just extensions of what Consolvo calls "gaming capital" (2007) or the information players naturally share with one another regarding how a game functions and the best strategies for playing.

Content Database Examples

Content databases, in general, attempt to provide players with extra information about a game but cannot necessarily be made entirely accessible in the game at all times. In other words, if a game could answer every question a player had about a game's content then there would be no need for content databases. However, topics like "best gameplay strategies" or "how to find an item" are not always easy questions answer. Discussing "best gameplay strategies", for example, may be hard because strategies within a game may often change. A patch for a game may be released that changes how items or characters work in the game. Even so, COD: Elite is one example play analytic

system that does try to offer basic information about gameplay strategy to its player population. Elite is built for Call of Duty: Modern Warfare 3 (COD:MW3), a first-person shooter game, and all of the strategies exist around attacking and killing other players in multiplayer matched. While using Elite, players can find a content database with gameplay strategies related to the maps, weapons, equipment and character classes that exist in COD:MW3. Maps include overlays describing how to attack different positions, while weapons and equipment are described based on their stats and the appropriate situations they should be used. The strategies provided are static, and may not make a player automatically a great player, but the content database at least provides players with a baseline of information they can refer to when trying to determine how to play the game.

Other content database examples follow with wiki model and a game's player community adds data to the system dynamically. WoWhead is a content database built for maintaining content related to World of Warcraft and allows players to upload/edit data. While WoW has a play analytic system with a content database, named the Armory, WoWhead provides additionally information that the Armory would never give out. For example, the Armory may have the stats of an item (damage, level, speed, etc.) found in WoW but a system like WoWhead includes data related to how often the item appears in the game. WoWhead obtains this type of data because the system has its own tracking system that players can download and use to upload data about their gameplay. By combining the data from various players an accurate value can be calculated and describe how often an item appears, for example how often the item appears on a certain type of enemy. Many enemies may "drop" the item, meaning the item appears on the enemy after it is killed by a player, and if the player needs to find a particular item then data regarding which enemy drops the item is valuable information. The Armory does not provide this type information to players because a) it can be argued giving out that type of data takes away from the exploration aspect of the game and b) if the developer says an item appears a certain amount of times, players may become angry if the item does not appear as often as advertised. Content databases are therefore products of how the owners of the

database wish to portray a game. Developers often want to provide more basic data to aid players during gameplay (but do not want to take away from the exploration aspect of the game) while player created databases, like WoWhead, provide highly detailed data in order to reveal as much about a game as possible.

Functional and Phenomenological Aspects of Content Database

Visible

Content databases make content made specifically for a game visible. Content related to the game environment and the game's mechanics often appear in these databases. This is so players can be given information about the game they are playing. COD:Elite, for instance, provides data related to: weapons, classes and map strategies. The Armory provides data related to content found in World of Warcraft, including: items, quests, NPCs, etc. Dynamic databases often go farther than static databases, revealing game content that is not normally provided by the developer. WoWhead built its own tracking addon for WoW to monitor how often loot drops from enemies and monitors where enemies spawn in the world. Both are facts about the world that are not revealed by the other WoW content database, the Armory. The Armory instead makes item and NPC data visible but does not give more information. The amount of data that is made visible is at the discretion of the system's creators.

Comparability

Content databases are perhaps the only play analytic category that does not seek to compare data together. These databases represent repositories of data that players can search and relate to their own gameplay. If a player notices an item found in Heroes of Newerth that they have never used they can search for and learn about the item in the game's content database. The information the player needs about the item may be in the game (e.g. the attributes or stats of items are provided while playing HoN) but the player does not have to start a game in order to view the data if they use the online database. Also, players can compare pieces of content on their own using a content database, if they choose. A HoN player can compare the attributes/stats of two items together to see

which item would benefit their play style but the items are not placed into direct competition with each other in the database itself as in other forms of play analytics such as global stats (e.g. usage stats do compare in-game items to one another). Content databases are the repositories that players use to gain extra information about a game's content that may otherwise be harder to access while playing the game and may use the information they find to compare how the content relates to their gameplay.

Control

Developers have control over a game's content and should have an easy time controlling a content database referring to that data as a result. Although, developers tend to only build static content databases as part of their play analytic systems, which may hide or not provide certain data points, and dynamic databases spring up in order to fill in the gaps. Take for example the two content databases built for World of Warcraft, the Armory (a developer controlled database) and WoWhead (a player controlled database). Both provide data related to in-game items, quests, character classes and NPCs. They both make it easier for players to search for data related to their gameplay without have to find the data as part of their gameplay experience. For instance, if a WoW player wishes to know the attributes of a specific item they have to find the item in the virtual world before the game provides the information. Using a database like the Armory or WoWhead a player can search and find the item without having direct access to the item in the game. The major difference between the Armory and WoWhead, however, is that WoWhead is a dynamic database.

WoWhead is constantly collecting data from players about the content found in WoW. Where the Armory provides item attributes, WoWhead provides: what enemies have the item, where those enemies are located and how much money an item has been worth on the in-game auction house over time. The data WoWhead provides is available in the game but it must be collected by players, otherwise the data gets lost as part of the game. For example, when a player kills an enemy that enemy drops certain items. If the player is using WoWhead's tracking program (an add-on that allows WoWhead to collect data when a player is playing WoW) the program sends the data back to the database

where it marks which items were dropped by the enemy. Normally this data could not be tracked outside of the servers running WoW but since WoWhead provides players with a means of tracking their data, the database is able to collect more data than is provided to players through the developer run Armory database. Some argue that collecting and provide access to so much data about a game is actually a form of cheating (Consalvo, 2007) because it becomes too easy for players to play the game; players consult the database whenever they are stuck and never learn to enjoy the exploration aspects of the game. This may be one reason why developer content databases rarely provide dynamic data to players. But there is no denying that databases like WoWhead exist and that these databases end up controlling, or at least exposing, more data related to gameplay than most of the developer produced content databases.

Exposure

Static databases expose all of a game's content in one place and make it easier for players to browse through the data, rather than trying to find the data within the game itself. For example, the World of Warcraft Armory allows a player to look up any item in the game. If a player wanted to find data about a particular item while playing WoW they would have to find the item in the environment or find a player with the item. Instead, by using the Armory any item can be viewed and this becomes particularly useful when dealing with games that have lots of items for players to choose from (as is the case with many MMO games).

Dynamic databases expose player knowledge in addition to exposing game content. Dynamic databases grow over time due to the constant influx of new data and information players provide. If a WoW player is looking for an item using the Armory (a static database) they are typically given the item attributes. If the player looks for an item in the dynamic database WoWhead players are given the item's attributes plus where they can find the item, how often particular NPCs 'drop' the item, how much the item is worth and comments submitted by players regarding the item (e.g. another player may post how the item works with a particular character ability or spell). This is not data that would normally be available to players but by collected data from multiple player

sources (using a tracking add-on that is run while a player is playing WoW) WoWhead is able to expose the extra data. Dynamic databases attempt to expose the data that players tend to share amongst each other (such as what is the best way to fight an enemy or where to find the best items) and end up exposing most of what a game has to offer. Static databases are built to expose data like a reference guide, providing standardized information that is useful but is devoid of personalization or experience.

Interconnected

Both static and dynamic databases connect players to a game's content. Content databases make it easy for players to find data they need to play the game and have, more or less, replaced the game manual. There are thousands of items, quests and NPCs in games like Lord of the Rings Online and World of Warcraft. Content databases help players navigate that data without forcing the player to hunt down the information in the game. Additionally, dynamic databases also connect players through 'gaming capital'. Consolvo defines gaming capital as knowledge players learn as they play a game (2007). The more a player tests and explores a game the more that player learns about the boundaries and patterns that exist in the game. The player gains gaming capital by acquiring knowledge about the patterns and boundaries of a game. In the past, players could share this knowledge with their friends but the knowledge would not expand further. With the creation of gaming magazines like Nintendo Power, gaming capital began to disseminate further than circles of friends through various forms of media. Tips and tricks for playing a game could be published to larger groups of players, making it easier for players that found themselves stuck to continue with their game. As content databases started appearing online they began taking the place of these other media sources and allow players to share their gaming capital with any other player using the same content database. Players on WoWhead, for example, can leave comments related to in-game quests and provide information on how to finish the quest quickly. Even though content databases mainly connect players to a game's content, they can also be used by players to share their experiences with a game's content too.

History

Content databases are built to expose game content and game content does not always change over time. An item added to World of Warcraft will likely be in the same six months from now. There are cases when expansions are released for games or patches alter the content of a game, but overall most of a game's content is solidified when the game is released. Thus, most of a game's content does not necessarily have a history to tell, the content exists and a database provides players access to the content. There are cases of dynamic databases attempting to track when new content is added. WoWhead, for instance, marks certain quests or NPCs as being added through one of WoW's multiple expansions (there have been three major content expansions for WoW thus far). Sometimes content was altered in-between expansions and WoWhead attempts to detail what changes occurred. But typically, there is little historical explanation of game content because content rarely has a history. Content is added to the game one day and remains unchanged until the developer makes an alteration, which is rare.

Content Database Summary

Content databases act as single source for data regarding a game's content. Static versions of these databases tend to be built by developers when their game contains many individual pieces of content and players need a source where they can reference those pieces of content. For instance, databases built for massively multiplayer online role playing games are a common occurrence (Aion, Lord of the Rings Online, World of Warcraft) because those games have many items, quests and other game content players need to reference. Although, other games such as Heroes of Newerth and COD: Modern Warfare 3 also have content databases containing data related to items and character classes.

Developers typically create static content databases because they do not want to give too much data away. A content database is seen as a reference guide to developers and players are expected to explore a game even though the database provides them with additional data about the game. Dynamic content databases take the opposite approach, these databases are built to expose as much data about a game's content as possible. These dynamic content databases are usually built by outside parties, like players, and

constantly collect data regarding a game's content and add it to the database. WoWhead is one example of a dynamic database made for World of Warcraft that continuously collects data from players in relation to where players find items in the game, where they defeat NPCs, where quests are located and how much items are worth on the in-game auction house (just to name a few). Most of this collected data is not released on the official WoW static content database (the Armory) because divulging such detailed data takes away from the exploration aspect of playing the game. There is very little developers can do about these dynamic content databases, however, and in some cases they eventually embrace them (WoW's Armory system started linking directly to WoWhead but only after both systems had been running for many years). Players find ways to share their knowledge about a game's content and dynamic content databases are one way play analytic systems allow players to do so.

Discussion

After reviewing over 80 different play analytic systems many of their similarities and differences emerge. It is quite common to find systems that combine many of the play analytic categories. For example, leaderboards can be used in conjunction with content generators such as the leaderboards Little Big Planet provides for each player created level. Maps are used by content databases like WoWhead to reveal where objects and entities are located in World of Warcraft. Group stats can aggregate each group member's player stats together to offer a collective set of stats, which is how the Battlelog system works for Battlefield 3. Game-related data is also shared between play analytic categories too. Player statistics and content databases can share data regarding items or NPCs found in the game (e.g. AION, Lord of the Rings Online and the WoW Armory). Global war features, as Brink or End War offer, collect data from individual player matches and use them to create a meta-level war players can take part in. The line between each category is fluid as many categories are combined within a play analytic system and the features/variation in these categories often share data.

The differences that exist between the play analytic categories revolve around what categories offer players compared to the other categories. One major difference is

whether a category offers player data related to their gameplay or a game's content. A player's gameplay factors into categories like leaderboard and global stats, and content databases provide players with data regarding the assets and components of a game. Gameplay data shows what players have done and content data shows players what is available in the game. Additionally, categories are differentiated by what experience they offer players. Gameplay categories like player statistics and group statistics can offer similar stats, even collect the same data, but the categories portray two different experiences: an individual's experience and a group's experience. The individual's data is their own, it is personal, and the group's data represents a collective effort. Player stats and group stats offer two different contexts for a similar set of data: one is for portraying a player's personal progression through a game and the other is for a group of players to share the progression together (and they possibly progress as single entity as is the case of integrated groups). Categories differ not only based on the type of data they expose (e.g. gameplay versus content) but also how the data is placed into context as part of a player's experience.

In addition to the similarities and differences found in this study, I have organized a list of common themes that relate to the overall design of the play analytic systems covered in this content analysis. I have split the discussion of these themes into the six functional and phenomenological properties in order to give a better sense of how the themes relate to the theoretical framework, and to show how the framework can be further used to critique play analytic systems. Further analysis or the implication of these themes are discussed in chapter seven and combined with the user study explained in chapter six.

Visible

Competition is the driving force behind what data is made visible in play analytic systems. Many play analytic systems are built for violent/war-based games and have a strong emphasis on players competing against each other, either literally (in the game) or figuratively (by comparing data using a play analytic system). There are many first-person shooter games (Battlefield 3, COD: Modern Warfare 3, Gears of War 3, Gotham City

Imposters, Halo: Reach, Killzone, etc.), strategy games (Demigod, Halo Wars, Heroes of Newerth, Starcraft 2) and other competitive games (Need for Speed: Hot Pursuit, Red Faction, SSX, World of Warcraft) with play analytic systems. One reason why so many play analytic systems are built for competitive, violent games is perhaps due to many of the big budget games developed are competitive and violent. Building and maintaining a play analytic system is still expensive (which is why Microsoft attempted to shut down the Halo Wars stat tracking system, see chapter one, and why so many other systems have been taken offline) and, therefore, the games that can support a play analytic system are the big budget games; those games that are competitive. Even content generators have ranking systems that create a sense of competition between players creating game content. Getting a higher ranking for a level created in Little Big Planet or creature created in Spore means more players are likely to see and use the content in their own game. Even the Sims Exchange, which allows players to create movies and write their own stories, has a “recommend it” feature where player created stories can earn more recommendations. Play analytics seem to be a perfect companion for visualizing competition. Play analytics makes players more aware of and able to refer to data used to compare a player’s skill level, their endurance capacity or their creative prowess; all instances of competition.

Not every play analytic system is focused on competition and many of the non-competitive systems are based around providing data to make a game easier or more accessible. Mapping systems like Minecraft X-Ray and Terraria Maps are built to help players find their way in either game’s respective environments by providing detailed maps. Similarly, other mapping systems like Batman Arkham City App or Map WoW plot the positions of specific items or locations in their games for players to use while they are in-game. There are also systems like save game editors (the save game editor for Torchlight, entitled Torchview, is one of the systems reviewed) that can both provide players with a way to edit their save game and a way to view their current stats and load-out. All of these non-competitive systems visualize data in order to make a game more accessible to players.

Looking for play analytic systems that are not based around competition or making a game accessible leaves very few systems available. Of the 80+ systems reviewed only one does not fit into the competitive or accessible category. Spore Skeleton visualizes the skeleton structures of creators created in the game Spore. The way in which the creatures are visualized is everything external to the creature – body shape, color and accessories – are removed and only the underlying skeleton is left. Spore skeleton does not make a creature's skeleton more accessible because players do not see or manipulate the skeleton directly while building creatures in Spore. A creature's skeleton is a by-product of the design experience, one that Spore Skeleton highlights. The creatures are also not compared as a means to incite competition as the system does not allow players to rank or manipulate the creature's meta-data (although only the top 100 rated Spore creatures are listed but this is only used as a means of ordering the models). Spore Skeleton is the only system reviewed that takes a non-normative approach, and a non-competitive/non-accessibility approach, to visualizing data.

Most play analytic systems are built for normative purposes with only a few examples covered in this study allowing for non-normative analytic approaches. Promoting competition or accessibility are the major normative factors that direct play analytics towards visualizing data for utilitarian and useful purposes. Only Spore Skeleton attempts to take a non-normative approach to play analytics and is reminiscent of artistic visualization discussed in chapter three. Spore Skeleton alters a viewer's perspective about their game data. Creature models are displayed in such a way in Spore Skeleton that they make use of the game-related data collected from the Spore API but only reveal the underlying data that is only useful to the game's graphics engine. Perhaps if more games offered access to their data, through APIs for example, more systems like Spore Skeleton would be created. Although, maybe the lack of non-normative play analytic system has to do with the fact that many games revolve around competition. This may force a design mindset that play analytics need to promote competitive behavior, or at the very least, make a game more accessible for players by visualizing useful data to aid in a player's progression through a game.

Comparability

Out of the five analysis categories described in chapter three (analytic, creation, monitoring, reflective and reminisce) monitoring, reminiscing and creation are the main analysis categories provided by play analytic systems. First, monitoring is the most prevalent analysis method found in play analytics. All variations of leaderboards are built around the concept of monitoring. Leaderboards keep track of each player's performance so players can compare themselves to, and compete with, other players. Activity feeds are used by player stats or group stats and allow players to monitor events of their fellow group members or friends. Content generators and content databases are built to search for data, to monitor what is available, and often highlight new data to players. Second, players are given the means of reminiscing about their data when a system records and displays past events (although I discuss the extent of which play analytic systems record temporal events in the history section below). Achievements, for example, are marked with a timestamp and give players a chance to reminisce about how, and when, the achievements were earned. Match stats also can provide detailed accounts of the events that transpired during individual matches and in some cases allow players to re-experience and reminisce about the matches (as is the case with replays). Third, some play analytic systems do allow players to create their own means of comparing data. APIs are offered for some games (Eve Online, Halo, Spore and World of Warcraft) and players have created systems like Spore Skeleton and WoWhead using those APIs. Content generators also allow players to create content and sometimes this content involves data related to gameplay. The Sims exchange, for example, has a story creation feature that gives players the ability to combine screenshots from their game and write stories as captions to the shots. There are also cases where players create a system for collecting their own gameplay data manually, for example the Darkfall Political Map system is built to crowd source data from players about the political alliances in the game. All possible forms of monitoring, reminiscing and creation analysis may not be represented by current play analytic systems (for example, not every play analytic system offers an API or stores historic data to facilitate reminiscing) but these three methods of comparing, or relating, data are the methods emphasized.

It is harder for players to analyze game data analytically or reflectively using play analytic systems because the data provided is often too general, aggregated or otherwise obscured from meaningful analysis. Player stats, for instance, often contain data collected over a player's entire time playing a game; making individual play sessions meaningless and hiding the fluctuation of their stat progression over time. It becomes a challenge to analyze the data both analytically, in order to find patterns in a player's behavior, and reflectively, in order to question why a specific set of data is collected or whether the data collected provides an accurate identity of the player. Group stats suffer from the same problem of using aggregated values, along with leaderboards and global stats. Content databases and content generators are regularly based around searching for values and do not provide methods for analyzing patterns amongst the entire set of game content. Match stats is the only play analytic category that routinely provides detailed data, with numerous dimensions, that can be adequately used in analytic or reflective analysis. However, there are not many match analysis tools that exist to analytically or reflectively analyze match data. SC2Gears and SC2Replay.com are the only two systems studied that allow players to analytically, and reflectively question, data collected from a match (in this case Starcraft 2 matches). Both of the systems provide players with a means of analyzing their data for behavioral patterns and trends in their Starcraft 2 gameplay. Both systems were also built to rival the match analysis support the developer of Starcraft 2 provides (which is severely limited in comparison to tools like SC2Gears), thereby reflectively questioning what data should be available to players. Outside of these two systems, however, analytic and reflective analysis is rarely afforded by play analytic systems.

In the visibility section above I mentioned that play analytics systems routinely visualize data related to promoting competition and accessibility. The third type of activity that is promoted by play analytics is sharing and it is caused by the emphasis play analytic system design places on the monitoring, reminiscing and creation analysis methods. Play analytic systems are built to compare players and for players to review data, not analytically or reflectively analyze data. Players can monitor their friend's activity feed or compare the number of achievements between themselves and their

friends but this is not analytic data analysis. Leaderboards and global stats often lack any temporal dimension to determine patterns over time and are regulated to the task of sharing the current state of a game instead. Even content generators and content databases use search as a means of sharing data between players but offer little in the way of finding trends amidst the content (for example, which players created the most content over the past week). There are often no means given for gaining insights while using play analytic systems, which could be used during future gameplay sessions. There is no ability to reflectively question what data is being collected or to organize the data into a new perspective. Lack of analytical or reflective analysis means play analytics are really used for sharing rather than the common interpretation of analytical data analysis, as the “science of analytical reasoning”, for forming hypotheses, gaining insights, finding patterns and providing recommendations.

Control

Many play analytic systems are tightly controlled by game developers. This is perhaps quite obvious because game developers create the means to collect data from the games they develop. Hence, developers have unbridled access to gameplay data and data related the content found in their games. Play analytic systems are easier for developers to build because they consistently have access to game-related data.

With developers keeping such a tight control over what data is available this may be one reason why play analytic systems do not offer as many analytical or reflective capabilities. Games are regularly designed around the principles of competitive play, accessibility and sharing. Play analytic systems reflect these design principles and many of the play analytic category variations described above reflect these design categories. Leaderboards, player stats, group stats, match stats and global stats promote competitive behavior and sharing data like scores, achievements and other quantitative values. Sharing also occurs using content generators as player create content for a game and then release it to the game’s player community. Content databases are built to provide a comprehensive catalog of a game’s content to make the content more assessable to players then would otherwise be available in the game itself. Since games are built to

provide competitive play, accessibility and sharing, only data related to those topics are made visible and compared using certain analysis methods (i.e. monitoring, reminiscing and creation). Competitive play is achieved through monitoring other player's gameplay. Players reminisce about data by sharing and monitoring data collected in the past. Finally, creation allows players to make the game more accessible by creating their own play analytic systems that have not been provided by a game developer, such as Map WoW or the Noby Noby Boy stat website. Since developers wish to design games with the principles of competitive play, accessibility and sharing in mind, the play analytic systems built for these games share the same design principles.

When control over game-related data is given to players, either through a developer giving players access to data or players collecting data themselves, they have created some of the few play analytic system that do allow for analytic and reflective analysis. Darkfall Political Maps, for instance, must collect data manually using crowd sourcing methods because the developer of Darkfall does not provide the adequate means of recognizing, globally, which player groups control which cities in the game. While the mapping system is typically used to create an up to date version of the current political standings in Darkfall, the system does keep an historic record of every change to the political climate in the game (which could be used to analytically analyze player group behavior over time). WoWhead, another player created play analytic system, also deputizes players as their data collectors by using an API modification for World of Warcraft. By collecting data from multiple players the system is able to analyze the data, create maps depicting where many of the in-game objects exist in the world (often these objects exist in more than one place) and can track how often specific events occur (such as when items appear after defeating enemies). Many of the analytic features of WoWhead were not provided by the play analytic system built by the World of Warcraft developers (although recently the WoW Armory has been linking directly to WoWhead). SC2gears, as I described in the comparability section above, is also a player created system that provides the means to analytically and reflectively analyze Starcraft 2 replays. SC2Gears has to literally deconstruct Starcraft 2 replay files in order to access the available match data because the files are compressed using Blizzard's, the developer

of Starcraft 2, own proprietary compression format. Finally, Spore Skeleton is a play analytic system that allows players to reflect on the composition of the Spore models created by players in the game but does not need to resort to odds means to collect data. The Spore API allows Spore Skeleton to directly access the Spore database that contains all of the models players have created for the game. Whether players have to take data by force or are given access to data directly from a game developer, they have produced an equally relevant, if not a more diverse, set of play analytic systems compared to game developers. It can be argued that because many games restrict or make it difficult for players to access game-related data there should be many more play analytic systems, especially those built by players, than currently exist.

Expose

Almost every play analytic system exposes player data publicly. I had very little trouble accessing the play analytics systems covered in this content analysis and very few had any access restrictions (e.g. Orcs Must Die does not allow players to share their player stats so the only way to see the player stat dossier is to purchase the game). Leaderboards and global stats seem to be the most accessible play analytic categories because they are typically general enough to be potentially relevant to everyone. General leaderboards list everyone's score and global stats are often used by a game developer to promote the size of their player community, hence why they are easy to access. Player stats, group stats, match stats, content generators and maps are usually placed behind an access wall, like a login system. These access walls are not hard to breach, virtually every play analytic system offers a free account option where players can join without even purchasing the game the play analytic system is built for. The main reason many of these play analytic categories exist behind a barrier of access is due to the system needing to know a player's identity. Player stats cannot be given to unknown players, hence group stats, maps, generated content and match stats (which are usually attached to specific players) must be paired with some sort of identification process like a login system in order to match a player to their data. But once a player has access, most gameplay related data is freely available for players to share and monitor.

Play analytic systems typically expose player data freely but game content data may be held back. As I have already argued, games (and play analytic systems) are regularly designed around competitive play, accessibility and sharing. However, these design principles can conflict. For example, too much accessibility could create a less competitive environment. It has been argued that when content databases provide too much accessibility, or access to game content, they can take away from the exploration and competitive aspects in games (Consalvo, 2007). WoWhead provides players with tons of data regarding how to find specific items/enemies and complete quests in World of Warcraft. So much so that it is fairly easy to progress through the game without having to explore the game environment (as a player would who was naturally trying to find items or complete a quest) or ask for help from other players. Systems like WoWhead and Map WoW provide greater accessibility to players, more so than the developer controlled WoW Armory, but that accessibility may take away from other aspects of gameplay such as exploration and competition. This is perhaps why developer produced content databases do not expose as much data related to in-game content because those systems may reduce the play experience.

Interconnected

With an emphasis on monitoring, creating, reminiscing and sharing, play analytics is very reliant on players connecting with one another and promotes a comrade mentality. Systems stress the ability to share data (for both competitive and reminiscing purposes) and identify oneself as part of the player community. It is incredibly easy to access statistics in most systems and content generators, in particular, are built specifically around sharing created content. Global stat variations like community events and global war are meant to bring a player community together to accomplish tasks that demand the entire community's attention. Friend leaderboards create sub-set leaderboards so every player can have a chance to compete against their friends and feel like they are progressing in the game. Group stats are based around players forming internal player communities that share stats and compete with, and sometimes against, each other. All of these examples point to an emphasis to allow players to connect to one another and to

form various levels of comradeship with different groups of players: whether with their friends, a larger group of players or the player community as a whole.

Many games do not have integrated social or communication features and players are forced to connect with other players outside of their gameplay experience. As a result, play analytic systems are acting as a means for players to connect and form groups related to specific games. By giving players a place to share data, compete and reminisce, play analytic systems act as hubs for players to communicate and meet each other. Additionally, play analytic systems are routinely connected to communication features like forums and messaging systems. Promoting a comrade mentality is perhaps the best thing play analytic designers can do in order to keep players engaged with a game and the play analytic system. By providing a hub for players to congregate, play analytics acts as an additional socializing experience that is connected to a player's gameplay experience.

History

Many of the play analytic systems lack features that present data historically. Data is regularly presented in an up-to-date format or continuously aggregated. Game content data is rarely shown changing over time in content databases, content generators and maps, where these categories give equal weight to two pieces of data – regardless of their temporal displacement – so long as they are still considered game content (e.g. a game patch may remove content). Leaderboards, player stats and global stats are typically aggregated over time and exhibit players in their current form, devoid of visible changes over time. Match stats do often add a temporal dimension to gameplay data but, as I have already argued, there are very few systems that take advantage of the data's temporality (SC2Gears being one of those exceptions). Having a lack of history may also be one reason why play analytic systems afford the ability to monitor data; it is an in-the-moment type of analysis where history does not have a major effect. The bias towards monitoring, communicating and sharing data – which often means monitoring, communicated and sharing novel or new data – may be one cause behind many play analytic system forgoing historical data collection (along with the cost to store data over time, but I explore this issue in chapter seven).

With few systems offering historical data, it is hard to analytically, or reflectively, analyze data along a temporal dimension. Of course, data analysis does not demand data have a temporal dimension (for example, analyzing game content could be valuable without a temporal component) but when referring to gameplay data, where gameplay happens across time, having temporal data can be important. Match stats is usually the play analytic category that connects time with gameplay data. Systems like Bungie.net, COD: Elite and Gotham City Imposters all offer kill and death events over time, plotting them on a map, and allow players to inspect each event. Although, each of these systems collect much more data but rarely present that data over time. SC2Gears gets around this limitation by tapping into Starcraft 2 replay files and gains access to all the data needed to recreate a replay of a match. However, games with replay features, or replay files, are rare. Also, without features like replays it is hard to connect event data to memories of a gameplay match. Achievements, for example, record the time when you earned the achievement but it is hard to reminisce about earning the achievement because there is no additional data tying the achievement to past gameplay. Achievements, and other play analytic features, do not work like lifelogging systems nor do they operate like a SenseCam (Gemmell et al., 2006) . In addition to it being harder to analytically or reflectively analyze gameplay data because of the lacking temporal dimension, it is also harder to reminisce about collected gameplay data if only a timestamp is used to provide context to players. If temporality was emphasized in play analytic systems more systems would likely stress analytic and reflective analysis, and reminiscing about past gameplay may become more prevalent.

Summary

Play analytic systems tend to be normative in their approach to making data visible (where data is presented in a utilitarian or useful fashion) and this approach is similar to the typical understanding of analytics (analytics being seen as normative practice, as I discuss in chapter 4). However, the overall capacity for players to analytically analyze, or reflect on, their data is rare. Other methods of analysis – namely monitoring, reminiscing and creation – are much more prevalent. The type of data

collected and the features provided tend to allow for those methods of analysis but do not go far enough in order to provide players with analytic or reflective insights about their gameplay. Knowing I have 100000 points less than the next person on a leaderboard does not help me change, or give me insight into, how I perform or question the methods of earning points. Play analytic systems, as they are typically designed, are useful for acknowledging competition, making a game more accessible and sharing data (often in relation to competition and accessibility).

As I argued in chapter one, analytics should not be considered solely a normative practice. Analytics should be allowed to be artistic, avant-garde, broken and frivolous. Although, it seems that while play analytics is normative – as a practice meant to make games more accessible, competitive and allow players to share data, i.e. it is a useful practice – it does not go as far as many analytic systems that offer analytical data analysis methods and provide analytic insight. A player can rarely understand how to increase their performance based on the analyzing features these systems provide. It is hard to find patterns and form hypothesis about the behavior of individual players or groups of players. The exceptions being systems like SC2Gears or WoWhead, and both offer some level of analytical data analysis methods to players. There are also few systems that allow for reflective analysis that “alter a viewer’s preconceptions” and take a non-normative approach to play analytics. Many systems are built on common preconceptions of games: competition, communication/sharing and general accessibility. Only the Spore Skeleton system did not fit the common mold, of the systems studied, and offers player a chance to reflect on game data.

Play analytic systems basically inhabit a middle ground between the normal definition of what analytics is, including what analytical data analysis provides, and the more radical version of analytics, non-normative analytics, I argued for in chapter one. Data is continuously exposed to players but the analysis features provided are not adequate to allow for detailed analytic data analysis. Data can also be hard to access, often lacks a temporal dimension and is regularly sparse in terms of what data is collected. This has made it harder for players to create their own play analytic systems,

although there are exceptions, but it also makes it harder for players, or outside parties, to produce non-normative analytic systems (broken, frivolous or artist systems). Two things that should be taken away from this content analysis of play analytic systems are 1) access to data needs to be streamlined and 2) there are many more ways play analytics can be designed to analyze game-related data than have been explored.

CHAPTER 6

PLAY ANALYTIC USER STUDY

Content analysis is used to study how play analytic systems function as part of the gameplay process but does not capture how players react to those play analytic systems. Gauging player reactions can be achieved using two types of study methods. First, a researcher can sit down with a player during a gameplay session, asking questions of the player as they use a play analytic system. This type of user testing and interview format gives direct access to the player's thought process and can provide more context as to how a player may actual use a play analytic system in reference to other games they play. The downsides to this type of testing are that a limited player population can be examined, the process is time consuming and players are not in their normal environment where they play games, which may cause them to behave differently. The second method for gauging player reaction is to use game metrics to capture telemetric events remotely while players use a play analytic system. In this case, players can interact with the play analytic system from anywhere they have access to an internet connected computer (or gaming system) and events from that interactions are captured for analysis. Remote testing allows more players to take part in the study, allowing them to participate at their convenience and collects data regarding quantitative events which can be compared between players. The downside to this method is that, similar to any categorization system, the captured events are only as accurate in so far as how they are defined and the context of what a player is doing when a specific event was captured may be hard to determine.

Both user testing and remote testing offer their own strengths and weaknesses but I have chosen to use remote testing as the method to gauge player's reactions to a play analytic system. Remote testing may not be able to capture everything about the player's experience, in a similar way a play analytic system cannot capture an entire game experience, but this form of testing allows more players to participate while capturing

events that can be directly compared for difference in play behavior. Captured events, i.e. game metrics, are used to compare a player's behavior to other players, as well as comparing a player's behavior to themselves. This gives a quantitative view of how a player is reacting to play analytics and can be used as a stepping stone to design further studies that make use of qualitative methods such as user testing.

The study I have designed using remote testing takes a player through a series of steps where each step gathers information about the player and their play behavior. The study consists of five steps which include both capturing events from gameplay and gathering answers from surveys presented to each player. The surveys act as a means to determine the player's perception of the study and their gameplay behavior. Answers from the surveys are correlated to the data captured during gameplay to gauge if the play analytic system being studied had any effects on the player's perception, in addition to the effects on their behavior which is determined by analyzing the gameplay metrics. Furthermore, the player population that participated in this study was split up into three groups. Each group received the same study but was given a different play analytic system. This was done in order to test the effects the three play analytic variations had on each participant's gameplay, given the group they were placed in. One group was given a system that displayed their own data and compared their data with gameplay data from a phantom player (a fake set of player data that was the same for every player participant in this group). The second group was shown only their own data and was not compared with any other player. The third group, acting as the control group, was shown no data at all but still asked to play the game again, as the other two groups were instructed to do as well.

This chapter covers the study data collected from the three groups of participants. First, the methodology and study design is covered, presenting each of the study's five steps. Next, the data collected from the main survey, gameplay and post-survey is analyzed for significance amongst the different participant groups. Finally, a discussion covering the study results and how they relate to play analytics is presented.

Methodology and Study Design

Every participant had to progress through five steps in order to complete the study: take the main survey, play the first game, view a player dossier, play the second game and take a post-survey. Descriptions of these five steps are provided below. Example survey questions, a description of the game played and the design of the player dossier presented to each group are provided.

Main Survey

Participants in the study begin with a main survey asking them a number of demographic questions are asked during the main survey. Some demographic questions are multiple choice while other questions allowed the player to write-in their answer. Examples include:

“How many hours do you play digital games in a typical week?”

“How much money do you spend on purchasing or renting games per month?”

“How old are you?”

All of the main survey questions have been used in a number of previous studies I have worked on with colleagues (Magerko et al. 2010, Heeter et al., 2011) and have been used in combination with analyzing remotely collected gameplay metrics.

First Gameplay Session

After completing the main survey each participant is asked to play the game “Do I Have A Right?” for a particular amount of time while their gameplay is recorded. “Do I Have A Right?” (Figure 6.1) is a Flash-based game, allowing the game to be played using most modern web-browsers with the Flash plugin. Do I Have A Right? (Dihar) is designed as an educational game to help players learn specific United States constitutional amendments. The player takes on the role of a lawyer who runs a law firm and is responsible for hiring lawyers who specialize in defending cases related to constitutional amendments. Each day, which represent levels in the game, clients of the law firm enter the play area and the player is responsible for matching each client with a

suitable lawyer based on the client's case. As the player matches clients to lawyers correctly they earn prestige points, meant to represent prestige the law firm is gaining by winning trials.

Prestige points, are used as currency in Dihar, allowing the player to hire more lawyers, upgrade office equipment and purchase advertising space in a fictional local newspaper. Prestige points can be lost if clients are not handled properly, such as matching them with a lawyer that does not know the constitutional amendment the client's case requires. Clients enter the law firm each day and must be matched with a lawyer, told to come back the next day or told they do not have a case (some clients have in-valid cases). After the day is over the player may spend any prestige points they earned on additional lawyers and upgrades.

Each participant is instructed to play Dihar for four full days before moving onto the next step of the study. Each day represented a "level" in Dihar and playing four days roughly equaled 10 minutes of gameplay time. The study system prevented the participant from moving forward in the study unless they finished four full days. Participants could play for longer but forcing players to play for, at least, four days means each player can be compared along a common temporal benchmark that each player reached.

Other than being asked to play Dihar for four days players were given no other information. Each participant's gameplay is recorded automatically without any involvement by the participant. The type of data recorded from a participant's gameplay included results from trials (a trial occurs after a client is matched with a lawyer), when players buy upgrades, how much time it takes to complete each day, etc. Once a participant reached the end of the fourth day they are allowed to proceed to the play analytic profile step.



Figure 6.1: Screenshot of ‘DO I Have A Right?’. Players hire lawyers with experience in US constitutional amendments and match clients with cases that relate to those amendments.

Play Analytic Dossier

After completing their first play session, players are funneled into three groups. Each group up until this current step received the same study. All the questions in the main survey are the same and the game itself (Dihar) is not altered for any specific group. The three groups, as stated above, are separated and given access to a different amount of data related to their play session during the last step. Any player data that is referred to in this step regards the action data collected about the player recorded in Dihar during the previous step.

Group one is provided with their player data detailing their actions from the last game, in addition to data detailing a fake player. Figure 6.2 and 6.3 shows the layout of the play analytic screen given to group one. Action data is displayed showing a player’s: trial cases wins and losses, prestige earned, successfully used amendments, lawyers they purchased and upgrades they purchased. Group one had an additional section at the

bottom of their play analytic dossier. A fake player's timeline data was presented along with a large message asking the player to compare their timeline with the fake player. This timeline compared the client interactions of a player (whether a player successfully/unsuccessfully finished a client's case and whether a client with a false case was turned away or not) and the upgrade purchases the clients made each day. The fake player data was taken from a previous play through I had personally played so the data represented a player who had previous knowledge of the game Dihar. Every player in group one received the same fake player data, which keeps each player's experience viewing other player's data the same. Players in group one are asked to take a moment to review their data and to compare their data to the fake player, but are not told that the fake player was not a real player.

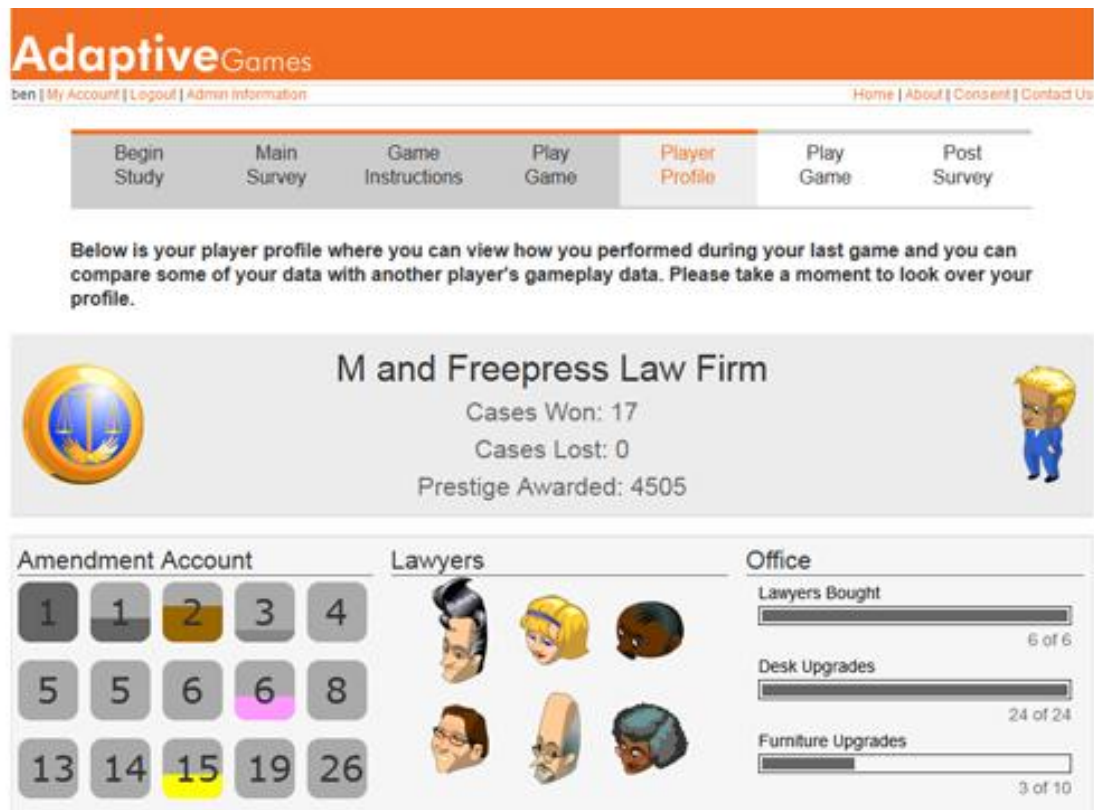
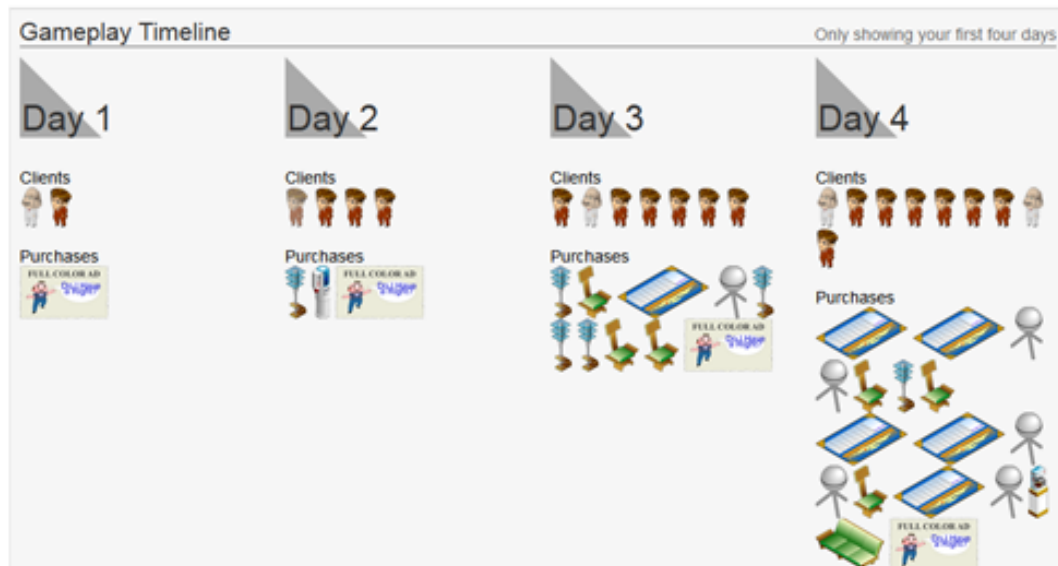
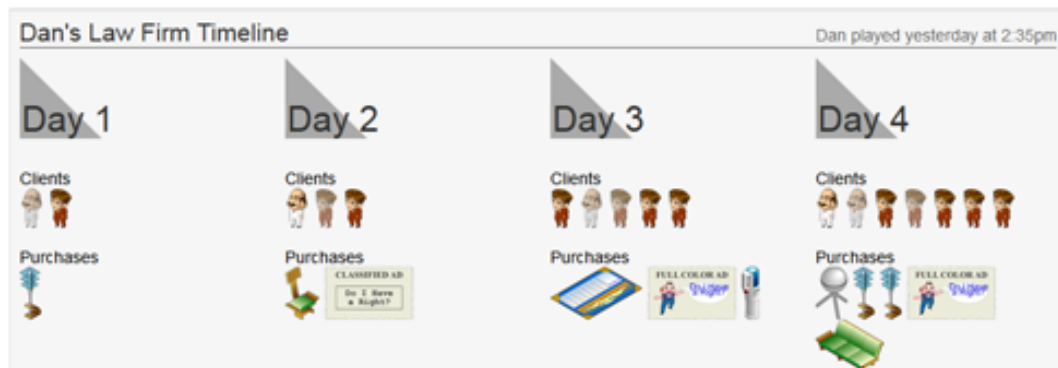


Figure 6.2: The top half of the play analytic dossier found in the game metric study.



Compare your Gameplay Timeline with Another Player



After you are done please click the continue button. Next you will play Do You Have A Right again.

[Continue](#)

Figure 6.3: The bottom half of the play analytic profile found in the game metric study.

Group two is provided with the same dossier layout as group one, however the fake player data is removed. Players in group two are asked, just as in group one, to spend some time reviewing their data before moving on. The differences between group one and group two are meant to gauge if comparing data with another player affects how a player behaves in their next play session. Finally, unlike the other two groups, group three is shown no data at all, representing the control group. They are instructed to move on in the study and play another game of Dihar. The other groups have to scroll down through their dossier before being able to click a button to continue the study. Differences

between the groups given a player dossier and the group given no data group are meant to provide acknowledgment of how player analytic systems alter a player's gameplay and the perception of their gameplay. A control group (group three) helps determine if factors such as performance, e.g. being able to earn more prestige points in this case, may increase as a result of simply learning how to play the game during the first play session instead of due to reviewing a player dossier.

Second Gameplay Session

After reviewing their profile, or in the case of group three being asked to move on to the next step, every player is asked to play another game of Do I Have A Right?. Nothing about the game changes from the first play session. In fact, everything is exactly the same right down to the type of player data recorded. Data collected during this session is used to compare to the player's actions from the first session. Based on which group the player is placed in, the differences between the two sessions provide a method of testing if displaying player data between play sessions affects player behavior. Every player is asked to play Dihar for another four days before being allowed to move onto the last step in the study.

Post-Survey

Players are not shown any player data after their second play session and are instead asked to fill out a final post-survey. Depending on a player's group each participant receives a slightly different post-survey. The differences between each post-survey involve asking groups one and two specific questions about the play analytic profile they saw in step three. Questions such as "Do you believe comparing your gameplay to another player in your profile was helpful? Why or why not?" or "Do you believe your player profile helped you earn more prestige points in the game the second time compared to the first?" are only given to groups one and two, and provide the means to figure out how those participants reacted to their dossiers.

Similar questions asked of every player include questions such as "Did you score more prestige points in your first or second game?". These questions test a player's perception of their gameplay across time. Groups one and two were given their total

prestige earned in the first game as part of their player profile which may help them remember those values compared to group three. These play session comparison questions help infer whether showing play data between rounds focuses a player's attention pushing them to monitor the data values, while they play, displayed in their profile. Additionally, all groups are asked questions regarding their experience with play analytic systems. These questions determine whether a player has dealt with play analytic systems before and to what extent, for example:

- “How often do you share photos or videos created in a game with other players online?”
- “Have you ever used an API provided by a game developer?”
- “Which of these games and their related data tracking services have you used?”

Upon completing the post-survey a player is marked as completing the study. All participants were offered extra credit as part of a university course once they completed the study.

Analysis of Study Data

The following data analysis consists of reviewing the data collected from the main survey, gameplay and post-survey portions of the study. Anova significance test were performed on all of the variables and survey questions discussed in order to find significance between groups. T-tests were used to determine if any significance existed between genders. Unless otherwise noted, no significance was found in the data presented either between groups or genders. The lack of significance found amidst the results could be caused by a number of issues, including: the low population size sampled for each group, the narrow age range of the participants, the recruitment of participants from a single university class, the education level of the participants and the choice to studying a smaller browser-based game. However, even though significance was not found in all cases, the analysis presented does identify some interesting trends that occurred, especially in regards to the differences between each group's gameplay behaviors. I speculate further in the sections below why some of the differences I mention may not

have been found significant in the study but how further studies need to be performed in order to determine the differences do not exist at all.

Main Survey Data and Study Population

The participants studied were young university students from a single large computer science related course. 60 student participants took part in the remote study with the median age of all the participants being 20 (18 being the lowest and 25 being the highest). 47 participants were males while 23 were female, with group one containing 14 males/6 females, group two containing 14 males/8 females and group three containing 9 males/9 females. On average, the participants had played a game the day they took the study with only one player stating they had last played a game more than a month ago. It is safe to assume that all of the participants had experience with games at the time of the study and most had recent experience.

Overall the median values of the participants regarding how often they play games were fairly moderate. Between all of the participants they played (based on the median) three games per week, spent somewhere between four to seven hours a week playing games (the median value being between '3 to 5 hours' and '6 to 10 hours'). Although, these values are only the median values and they hide the extremes that exist amongst the participants. For example, 12 participants said they played more than 10 hours a week and all of those participants were males. Only six females said they played games '6 to 10 hours' a week versus another 12 males who said the same (bringing the total to 24 males who played more than 6 hours per week). There was a significance of .0001 between the amount of time the female participants played games per week (with a mean between '1 to 2 hours' and '3 to 5 hours') compared to the male participants (with a mean between '3 to 5 hours' and '6 to 10 hours'). There was also a significance of .1 between male and females regarding playing web browser-based games, with females playing '1 to 2 hours' on average and males playing 'less than an hour'.

Participants were also asked how much money they spent on games each month. Participants spent less than \$15 on games per week and spent nothing on downloadable

content. Also, unlike the extreme significance between the time male and female spend playing games, the median value of how much money a participant spent on games on games per month is fairly accurate when looking at the extremes. Only four participants said they spend more than \$60, but less than \$120, per week on games. 23 participants said they spend no money on games at all and only one participant said they spend more money on DLC than games. It is safe to assume that many of the participants are not buying new, \$60 retail price, games often but do have access to many games (for example, free to play games, browser based games, mobile games or have access to friends with games). The lack of game purchases could be due to the study population consisting of university students who may not have a lot of disposable income.

Participants were also asked to rate their appreciation for various game genres as part of the main survey. Of the genres listed many were found favorable among the participants, include: action, adventure, brain games, first-person shooters, massively multiplayer online games, role-playing games and strategy. The lowest rated genres among the participants were exercise and sports games. There was also significance (with at least a value of .05 and some with .0001) found between the ratings the male and female participants gave to certain genres. Females rated brain, exercise, learning and puzzle games higher than males, while males rated action games, first-person shooters and role-playing games higher than females.

Gameplay Data

Gameplay data was captured from each participant as they played 'Do I Have A Right?'. This included capturing data related to earning prestige points, hiring lawyers, and counting the number of clients who enter the participant's office (just to name a few). Variables representing each player's gameplay data were summed up as totals or counted as frequencies; for example, the variable 'Cases Won' refers to the frequency (i.e. number of times) a player won a case in the game. Using these variable values, two types of analysis were used in order to determine differences in how players were behaving between their two play-throughs. First, significance tests were used to determine if a participant's gender or group affiliation were related to the variable difference values (i.e.

how much change occurred from a player's first game to their second game). The difference values between the following gameplay variables were calculated for each participant: cases won, cases lost, upgrade desk, upgrade office amenity, buy an ad, earned prestige points, and lost prestige points. No significance was found when comparing these difference values to gender or study group (even when group one and two were combined). However, there are mild differences between the separate player groups that should be noted and might lead to further investigation using a larger study group size. This led to the second type of analysis. The median values of each player group's variables are compared to find 'average' differences that exist between groups. A few exceptions exist between the group medians and the findings are discussed below.

The findings from the gameplay analysis below relate to play analytics and the player dossier system used by group one and two. For instance, I provide data related to how many players in each group bought more or less upgrades in their second game versus their first. Both upgrades and clients were shown as part of dossier given to group one and two (Figure 6.2 and 6.3). Some of these findings are also explored further by reviewing variable values within each individual group's data. Looking at individual player data helps explain the median 'averages' values and point to more nuanced behavior that would otherwise go unnoticed when only analyzing average values. Below I break up my analysis into smaller sub-sections; each sub-section has a title related to the variable values being analyzed.

Time Played

During the study, players were asked to play 'Do I Have A Right?' until they reached the end of the fourth day in the game. Days represent levels in Dihar where more clients come into the player's law office each day. While players were allowed to continue with the study after the fourth day many players continued to play the game after the day was completed (the game ends after the seventh day). Group one participants, particularly, both were capable of completing the first four days quicker than the other groups (finishing, on average, a full two to three minutes before the other group participants) and spent the most time playing the game after the four days were over.

Both group one and two spent roughly the same amount of total time playing the first game (many players finishing the full seven days of the game) but in the second game group two's total time greatly dropped off, matching group three's level of commitment. Group one's total time dropped between the first and second games as well but compared to the median amount of time it took players to finish the first four days, group one players still played an additional three minutes of time past the fourth day, compared to less than a minute other group participants played.

Table 6.1: All values in seconds except percentages. Difference percentages calculated by dividing difference values by the 'Time to Complete Day 4'

Game 1 Data	Group 1	Group 2	Group 3
Total Time Played	(seconds) 1746	1990	1617
Time to Complete 4 Days	(seconds) 984	1136	1223
Difference	(seconds) 762	854	394
Difference Percentage	77%	75%	32%
Game 2			
Total Time Played	813	763	685
Time to Complete 4 Days	618	710	633
Difference	195	53	52
Difference Percentage	31%	7%	7%

Group one was the only group given access to another player's gameplay in their provide dossier. As I explore in other sections below, having access to another player's gameplay may have provided players with additional perspectives on how to play Dihar. This may have helped keep group one participants engaged for longer compared to the other groups. However, given that group one participants were the quickest players to reach day four in the first game (while also spending more time playing the game after the fourth day), being equally as quick in the second game may just be a result of the players enjoying the game more so than the other group participants.

Clients

Connecting clients to lawyers is the main action players have perform in order to earn prestige points in Dihar. As more clients enter a participant's office they could potentially earn more points for connecting clients with one of their lawyers.

Additionally, when clients entered the participant's office in the game players had to decide whether or not the client had a valid case. A valid case meant the client had a situation that could be resolved by one of the constitutional amendments available in the game (for instance, a client may have a case where their first amendment right to 'freedom of speech' was revoked). The number of clients entering a participants office and the number of times a participant correctly verified a valid client case were both recorded during a participants gameplay because they both affect how many prestige points a participant earned.

Three variables related to the number of clients that entered a participant's office and the player's ability to verify a client's case were analyzed across the three groups. The median values of these variables are calculated across each group and separated by the two games the participants within each group were asked to play. I discuss the differences between the median value from the first and second games, within each group, below.

The first variable analyzed between the two games was the difference between how many clients a player saw between the first and second game. 45% of players in group one and two had an increase in the number of clients they saw in their game, compared to only 38% players in group three. This increase in clients may help explain why players in group one and two had more prestige points in their second game, as I discuss in the next section entitled 'Cases'.

The second variable analyzed (between each player's game one and two) was whether a player made fewer valid client rejects. A 'valid client reject' means the player rejected a client, saying they had no case, when in fact the client did have a valid case relating to one of the constitutional amendments. If a player rejected a valid client they lost as many prestige points as they would have earned from winning a valid case, making the act of rejecting valid clients expensive. 35% of group one players were able to decrease the number of rejected valid clients in game two while 45% of group two

player decreased their rejected valid client totals. Only 27% of group three players were able to decrease their rejected valid client totals.

The third, and final, difference variable flips the second variable around. Instead of looking at rejected valid clients, the third variable looks at accepted in-valid clients. Some clients in the game have in-valid cases and their case does not relate to any constitutional amendment. If players did not turn that client away then they risked losing prestige points when they lost the client's trial. 35% of group one players and 27% of group two players were able to decrease their accepted in-valid client numbers. 56% of group three players, however, were able to decrease their accepted in-valid client numbers.

Group one seemed to be more willing to accept any client that came into their office, during the second game, whether they had a valid case or not. This would explain why they had less 'valid client rejects' but more 'in-valid client accepts', because they were not turning away clients. An increase in the group one participant's willingness to accept clients could have been a result of viewing another player's data as part of the dossier system. Comparing clients was one of the visual comparisons provided by the dossier (Figure 6.2 and 6.3), which would draw a participant's attention, and in the Post-Survey below I offer reports of players mentioning they felt they needed to do better (earn more prestige) after comparing their data to another player. Comparing the number of clients using the dossier system may have made group one participants more accepting of clients because accepting more clients meant a higher probability of earning more prestige points (which group one did earn more prestige points, on average, during the second game when compared to the other groups, which is explained next, so being more accepting did not ultimately hurt them).

Cases

Winning and losing client cases in Dihar are the main game events that earn and remove prestige points from a player's prestige total. Four variables were analyzed relating to the outcome of cases and prestige points earned. The first two variables take

the median number of cases won and lost for each group. Group one participants were the only group to win more cases – moving from 6.5 cases to 8 cases- and lose less case – moving from 1 lost case to 0 - between the first and second game. Group two players were able to win slightly more cases – moving from 7.5 wins to 8 wins – and remained at 0 lost cases, on average, between both games. Group three’s median win/loss variables did not change, staying firm at 7 wins and 1 lose.

The third and fourth variables analyzed were the amount of prestige earned by the player and the amount of prestige lost. Similar to the win/loss variables, group one did better than both group two and three when it came to earning more prestige and losing less prestige between games. Group one players earned 278 more points of prestige in their second game versus group two’s 105 points, while group three players, on average, earned 128 fewer prestige points. Group one also lost 100 fewer prestige points to wrong decisions, with group two losing only 5 points less and group three losing 20 points less. (It should be noted that in Dihar if the player’s prestige score is at zero points, any event that causes the player to lose prestige points does not subtract those points, which would give the player a negative value. The prestige lost variable described here sums all the prestige loss events whether the player’s score reflected the subtraction or not. The reason being, these variables show how many prestige points a player potentially lost due to wrong decisions. Any decrease in the amount of wrong decisions, whether they affected the player’s score or not, point to an increase in the participant’s performance performance.)

Comparing group one participants to another player using the dossier may have helped push those participants to earn a higher score. They were more accepting of clients, as I explain in the last section, so they had a higher potential of earning more points. Some group one participants also later, in the post-survey, stated that the dossier made them realize they needed to do better and this may had an effect on how focused those participants were in regards to earning more points.

There was no significance found between each group but the fact that group one was the only group to improve, even if only slightly, may mean the dossier report did help players focus on their performance. Also, the differences between each group may be small, and not statistically significant, because Dihar is a smaller game. Most players are able to finish the game in less than 20 minutes and the range of “different” player behavior may be smaller due to the length of the game. If a different game was studied, and players were monitored over a longer period of time, the differences between the groups may have been more pronounced.

Table 6.2: All values are numerical points found in ‘Do I Have A Right?’ and represent the number of prestige points earned by participants over the first four days of Dihar gameplay.

Game 1 Data	Group 1	Group 2	Group 3
Prestige Earned	1922	2000	2000
Prestige Lost	200	205	220
Game 2 Data			
Prestige Earned	2200	2105	1875
Prestige Lost	100	200	200

Upgrades

Earning prestige points in Dihar means a player has more points to spend on buying upgrades for their office. Upgrades can be made in one of three areas: 1) lawyer desks, making lawyers more efficient at handling cases; 2) office space, makes the waiting area of the office more hospitable for clients (keeps the clients from becoming angry while waiting); and 3) buying advertisements, different advertisements bring in more clients the day they are bought. All three types of upgrades were tracked and the differences between how many of each upgrade a player bought were analyzed. Desk upgrades dropped from 1 to 0 for both group one and three, while holding at 1 for group two. Group one did have an increase in both the office space (from 1.5 to 2) and ad upgrades (from 0.5 to 1). Group two’s office space and ad upgrades dropped while group 3’s office space and ad upgrades stayed the same.

Another analysis done relating to upgrades was tracking how many participants altered their upgrade strategy overall. A change in strategy is defined as the player who increased or decreased the number of upgrades they bought by two items. If a player bought one more desk upgrade and one less office space upgrade they are considered to have changed their strategy. 41% of group two players changed their upgrade strategy from game one to two, while group three had a similar figure of 45%. However, 70% of group one players changed their upgrade strategy between games one and two.

Upgrades, along with clients, were the second type of data used to compare players together in the dossier system given to group one. Group two was also shown, graphically, how many upgrades they had bought but this did not seem to force them to re-think their upgrade strategy. Group one, however, did change their upgrade strategy, buying more office upgrades and advertisements as a result. Comparing group one participants, across clients and upgrades, to another player seems to have had at least a small effect on how they played their second game.

Post-survey Data

After completing the first survey and playing both games, players were asked to answer a final post-survey. The survey questions found in the post-survey asked players direct question in relation to play analytics. These questions were not asked during the main survey because if participants knew the study was about play analytics they may have altered their gameplay to be more willing to accept the dossier system. Below I have combined the participant answers into different sub-groups, in the same way I did for the gameplay analysis. Each of these sub-sections discuss a topic either directly related to play analytics (for example, participants were asked which play analytic systems they use) or related to other topics connected to play analytic (such as asking if players manage a group of players). Again, no significance was found between groups or genders unless noted in each sub-section's analysis.

Releasing Data

As I argued in chapter five many play analytic systems publically provide player data, for both users of the system and outsiders. Three questions in the post-survey were directed at determining the participant's feelings about releasing their gameplay data. When participants were asked if they would allow us, the researchers, to provide their gameplay data to other players every participant agreed to disseminate their data. When asked if they would be comfortable with us releasing their gameplay data with their real name, only 55% of the participants said they would be comfortable releasing their real names. Finally, when participants were asked if their data would be helpful to another player, 76% of participants thought their data would be helpful. Participants, therefore, were willing to share their data (so long as important personal information is kept safe) and felt that their data would be useful for other players, making the case that play analytic system work because players want to share their data.

Game-related Data

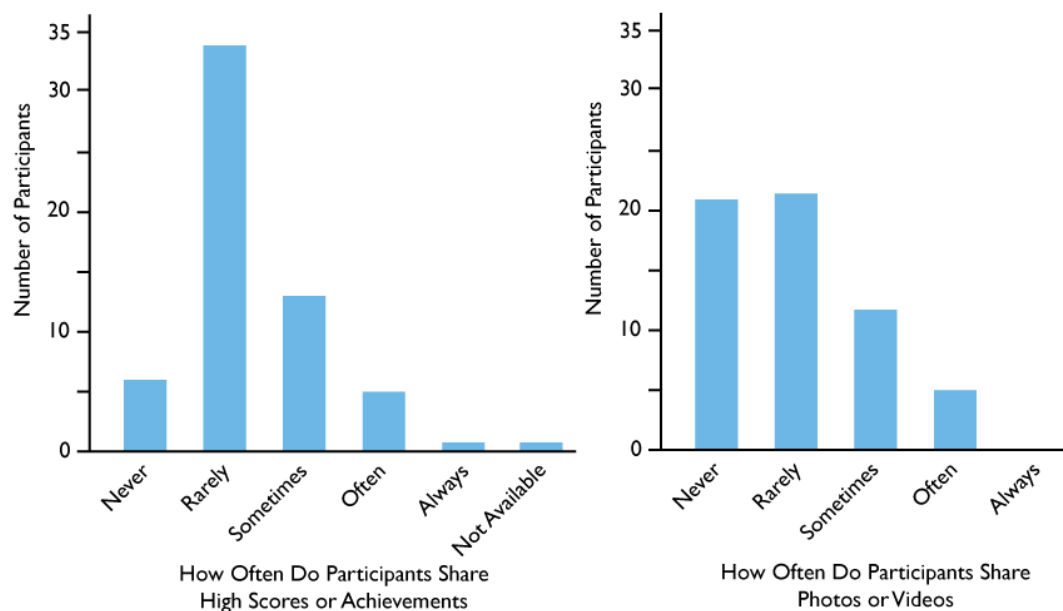


Figure 6.4: Graphs showing how often participants share data related to high scores/achievements (left) and photos/videos (right).

Another aspect of play analytic systems discussed in chapter five is the amount of game-related data provided to players (e.g. player gameplay data) and the amount of data players create (e.g. textures or models). Four questions in the post-survey related to determining the participants relationship with sharing and creating game-related data.

The first two questions ask participants whether they share high scores/achievements they earn in a game and if they share screenshots/videos they capture while playing a game (high scores, achievements, screenshots and videos being common game-related data found in play analytic systems). Participants were asked to answer based on a five part scale: never, rarely, sometimes, often or always. Most participants said they rarely shared data such as high scores and videos (with the median value being rarely for both questions).

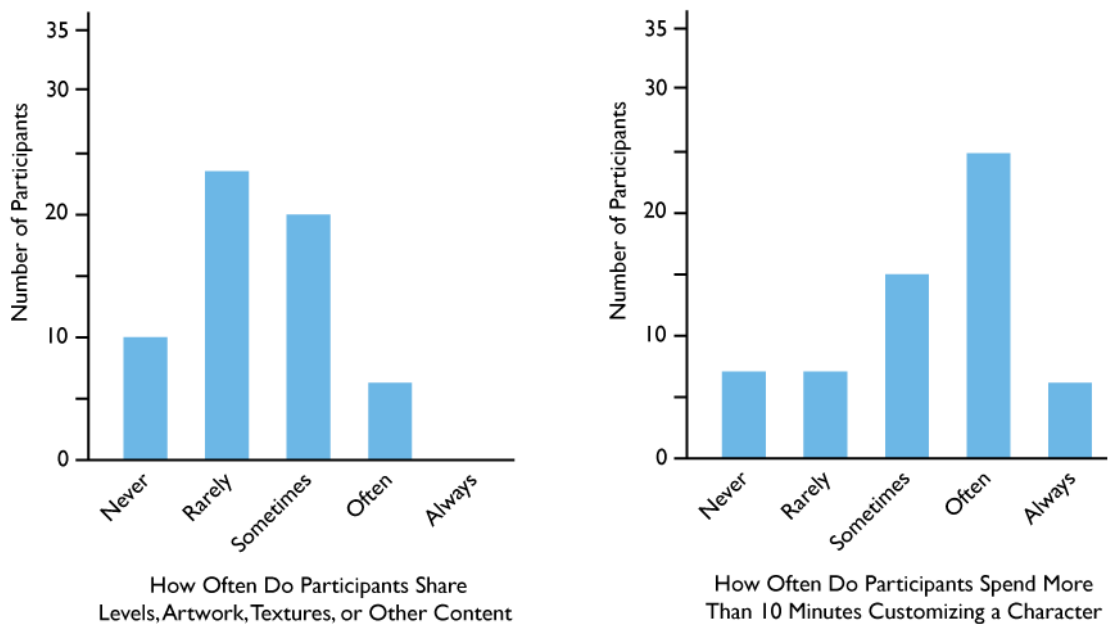


Figure 6.5: The graph on the left shows how often participants share data related to levels, artwork, textures and other content. The graph on the right shows how often participants spend more than ten minutes customizing their in-game character if the game allows character customization.

The third question asked participants concerned whether they created levels, artwork, textures or other game content when playing games using in-game editors. Participants used the same five part scale (never, rarely, sometimes, often or always) as the first two questions. Again, participants said, on average, they rarely created game content but more participants said they ‘sometimes’ created content compared to sharing content such as high scores or screenshots. The fourth question targeted a specific type of content creation in games, a player’s avatar. Many games allow players to alter their avatar’s appearance using an avatar editor (in a similar way games provide editors for producing levels, textures, etc.). The study participants on average said they ‘often’

edited their character's appearance for more than ten minutes. It seemed as the creation of data became more relevant to a player's personal gameplay, and often represented something the player used in the game, they were more willing to spend time creating the data. Perhaps play analytic systems should focus on visualizing data that represent gameplay objects and connect those objects to the data recorded from a player's gameplay rather than visualize abstract values such as scores or frequencies.

Altering Game Data

Besides players sharing and creating game-related content within the context of play analytic systems, there are also external means of producing game-related content. APIs, for one thing, can be used by players to create their own play analytic systems. APIs allow players to communicate directly with a database controlled by another party and grant access to game-related data that party owns (which is often a game developer controlled database containing player gameplay data). Some games also allow players to create MODs (modifications) that alter the appearance of the game (e.g. new artwork), add additional content to the game (new character models) or how a game functions (new game mechanics). Participants in this study were asked whether they had used APIs or ever built a MOD, to see how knowledgeable the participants were regarding these concepts.

About half the participants said they had used an API before (53%) and the other half had not. Only 41% of players had ever created a MOD for a game. Many of the players who had created a MOD were also male and a .01 significance was found between the genders. Group two also had more 'yes' answers to both questions than other groups. It seems group two had more participants that went beyond the features and services a game provided. If we include analysis from other questions, group two had more players that mod games, use APIs, spend time altering their character and creating level/artwork/other content. Since the participants were students at a technical university they may have been more willing to try modding games or use APIs. Further research would have to be done to see if large portions of a game's player population make use of APIs or create MODs.

In-game Information Formats

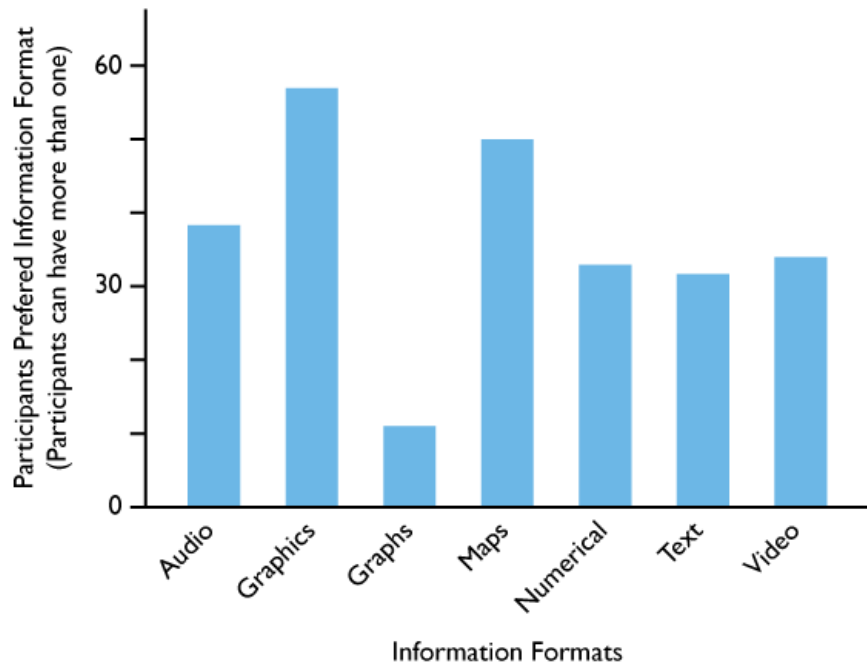


Figure 6.6: A graph showing how many participants preferred a certain information format when they were asked which formats are the best for conveying information during gameplay.

One question the study participants were asked concerned what type of formats they preferred when receiving information from games. This question was tangentially related to play analytics because receiving information in a game is different from receiving data outside of a game using a play analytic systems. However, both in-game and out-of-game systems use similar information formats. These formats include: audio, graphics, graphs, maps, numbers, text, videos. When participants were given the option to choose any number of formats they prefer, they chose graphics and maps as the most common format they prefer, with audio, numerical, text and video being preferred next. Graphics often appear on HUDs (head-up displays) in games (a common feature found in many games) and maps are also used regularly in games. Audio, Numerical, Text and

Video are common elements in games (often used to provide information) but are not always used while playing because they can clutter the screen. For example, video would not be used while a player is in an intense situation (where reflexes were required) but would be used in-between tense moments, cut-scenes for example. Audio is an exception to the “get in the way” argument, but audio cues often only provide information about one thing at a time, making it hard to use audio as a continuous information source (e.g. it is hard to hear multiple sounds at once, or only one person can talk at a time). Graphs were rarely chosen by the participants. The question asks which format the participant prefers “while playing a game” so viewing a graph while playing a game does not seem to be the best way to provide information. As to whether graphs provide valid information outside of gameplay, in a play analytic system for instance, is another matter but the fact that players prefer these other formats while playing games may be a sign that play analytic systems should adopt these formats further.

Manage Groups

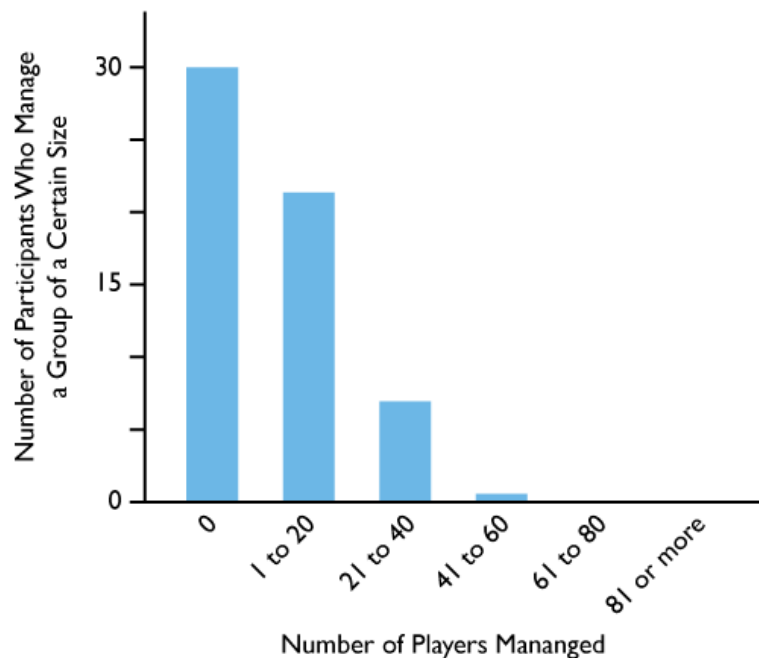


Figure 6.7: A graph showing how many players a participant manages as part of a player group, guild or clan.

In order to get a sense of how social and involved a participant was with organized player groups, each participant was asked if they managed a group of players. As discussed in chapter five, many players form groups for games (which is one reason why group stats appear in many play analytic systems). Of the participants that took part in the study half of them did not manage a group, with the next highest group managing one to twenty players. One participant marked they managed a group of 40 to 60 players.

Spectating

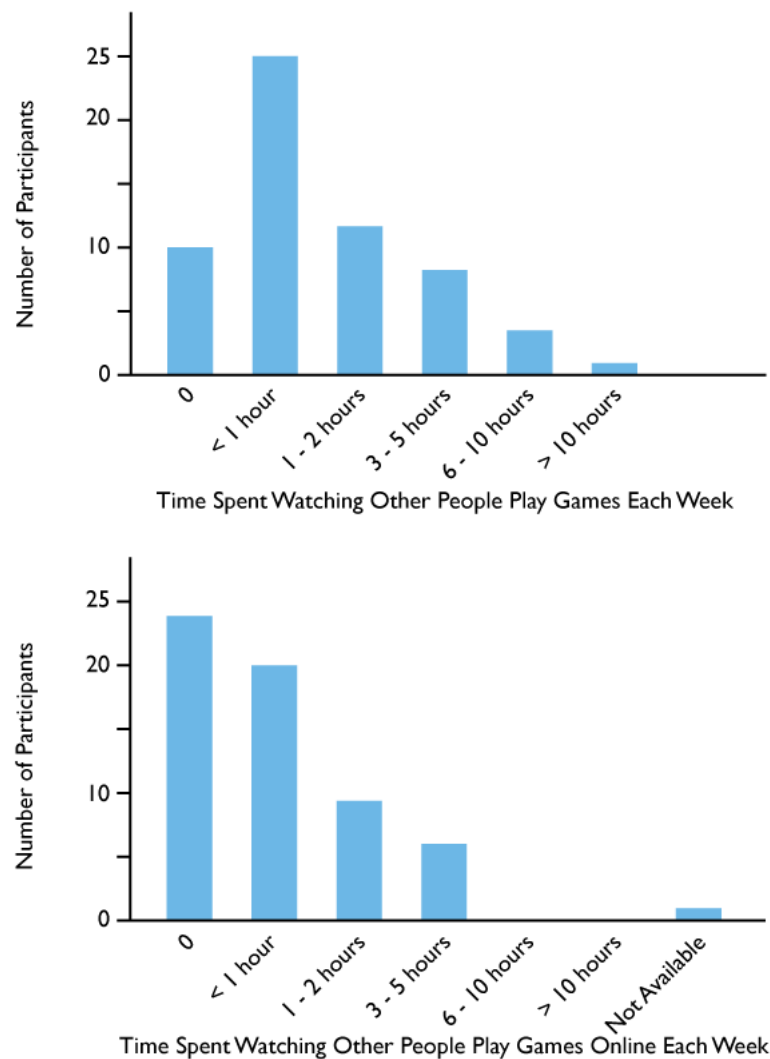


Figure 6.8: Graphs showing how often participants watch other people play games per week (top) and how often they watch other people play games online per week (bottom).

Spectating is related to watching/analyzing replays and following other players, both being features found in play analytic systems. Two questions were posed to each participant to determine how often they spectate other players: how often they watched other players play games while physically in the room with those players and how often participants watch other players play games online. Study participants watch other players they are physically in the presence of, generally, for less than two hours a week. Although, some participants marked that they watched other people play games for up to 10 hours a week (with one player watching 10+ hours a week). Spectating games while physically present seems to happen at least once a week for the study participants. There was also a significance of .05 when comparing female and males in the study participants, male participants spent more time watching other players than females.

Participants spent less time watching gameplay online than compared to watching gameplay in the same location. 'Less than an hour' was the median answer and no participant said they spent more than 5 hours a week watching gameplay online. This may mean spectating a game in person is easier to approach since the gameplay is perhaps happening in a communal areas (like a living room) and spectating gameplay in person may be a means of socializing. Watching gameplay online means someone must make a choice to seek out and watch the gameplay (although it is perhaps a more solitary, or personal experience). This may have implications for play analytic systems. 'Companion applications', systems that link to a game but are used by spectators, may be an interesting avenue to take play analytic systems. There may also be possibilities to expand on the type of spectating features certain games and play analytic systems offer (this is discuss further in the discussion below and in the dissertation conclusion).

Web Browsing and Gaming

Play analytic systems exist online and therefore participants were asked whether they spent time browsing the web while playing a game. The distribution of the participant's answers is fairly uniform along a scale from 'never' to 'always', with the median answer being 'sometimes'. Using a web browser while playing a game is not out of the question, it seems. Another question related to using the web to find frequently

asked questions, walkthroughs or cheat codes was also asked of each participants. Again, the median value found was ‘sometimes’ but more players leaned towards ‘rarely’ than they did in the first web browsing question. This means players may be willing to use play analytic systems while playing a game, in addition to using a system when they are not playing.

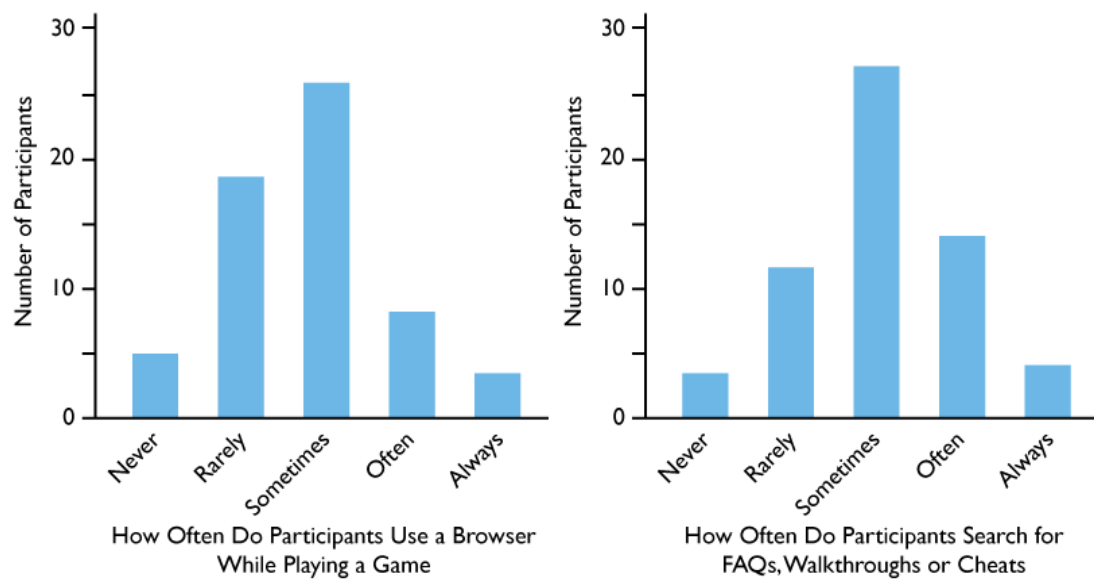


Figure 6.9: On the left, a graph showing how often participants use a browser while playing a game. On the right, a graph showing how often participants look up FAQs, walkthroughs or cheats for games.

Play Analytics

Participants were asked directly whether they used certain play analytic or game communication systems. Many of the larger play analytic systems discussed in chapter 5 were listed for players to choose and participants were able to choose any of the systems they used. The Assassin’s Creed play analytic system was the most popular system marked. The play analytic systems built for the first-person shooter games listed also had large followings, this included: Bad Company 2, Battlefield 3 (Battlelog), Call of Duty: Modern Warfare 3 (Elite). Other shooters – Killzone series, Resistance 3 and Gears of War 3 – were marked less often, however. Games in other genres had moderate followings too: League of Legends, The Sims Series, Warcraft, Starcraft Series (Battle.net), World of Warcraft (Armory).

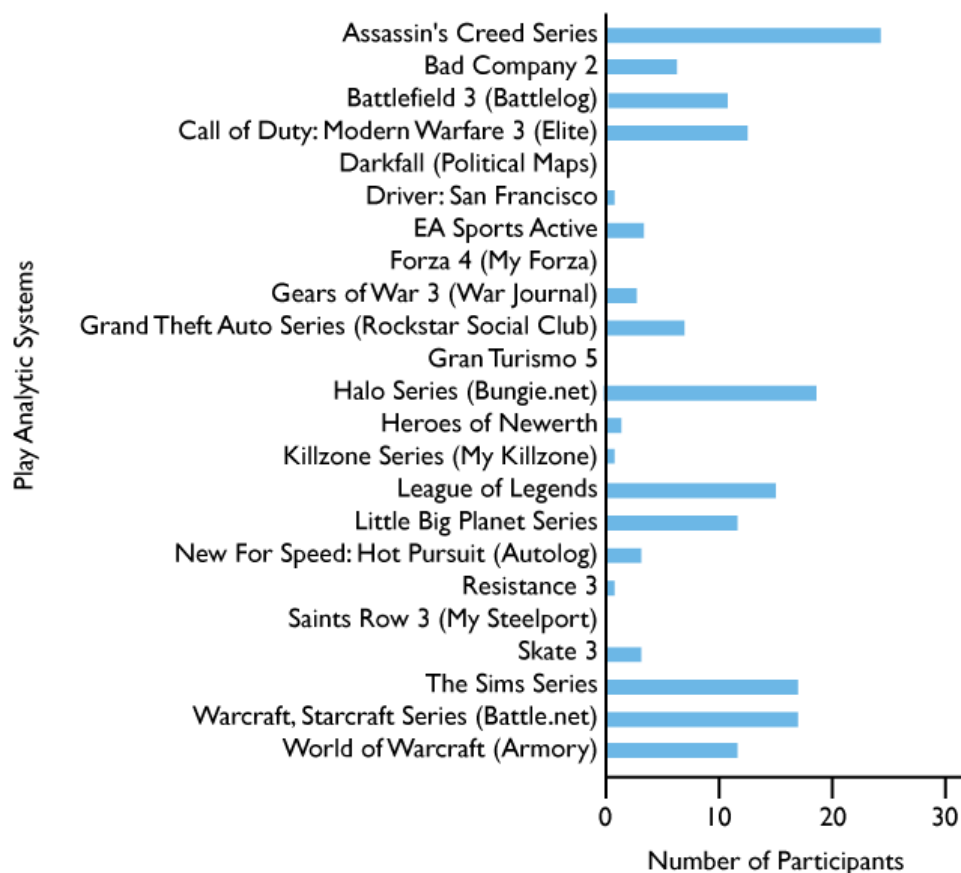


Figure 6.10: A graph showing how many participants used specific play analytic systems and played their games.

Participants were also asked whether they used certain game communication systems. These systems represent platforms players can use to communicate about games and has some relation to playing games. Some of the systems listed are not built solely for games, like Facebook or Google+, but others systems are dedicated to gaming audiences, like Raptr. These systems also included gaming platforms like Playstation Network, Steam and Xbox Live. Overall, the two major social networks were the most popular (Facebook and Google+) with the gaming platforms being the second most popular (Playstation Network, Steam, Xbox Live). The Flash gaming platforms also had at least a few participants. The rest of the communication systems had fewer, or no, participants using them.

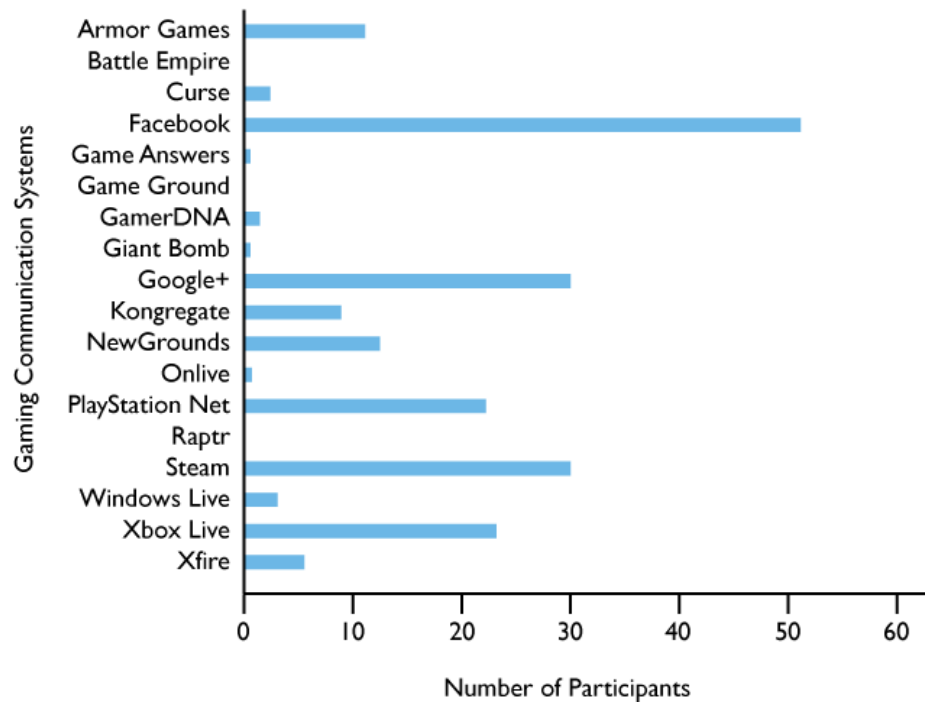


Figure 6.11: A graph showing how many participants use specific game-related communication systems.

Experiencing Play Analytics

As part of the post-survey, group one and group two were asked to answer questions regarding their experience using the player dossier system that appeared in-between their two gameplay sessions. These questions fell under three general areas of inquiry: 1) did play analytics affect a participant's performance or push them to try other strategies, 2) can play analytics help participants remember their gameplay, and 3) how did participants perceive the player dossier they were shown during the study. Each of these areas of inquiry are discussed below.

Play Analytic Effects on Player Performance or Gameplay Strategy

Participants were asked two questions regarding their gameplay performance and strategies. First, in response to a question asking participants if the player dossier provided help them earn more points in the second game, 40% of the participants said yes (17 out of 42). When asked if the player dossier pushed them to try a different strategy 52% said yes (22 out of 42). Comparatively, 12 out 20 participants in group one said yes

they tried a different strategy because of the player dossier, while only 10 out of 22 participants in group two said the same. Group one participants may have been primed to try a different strategy, more so than group two participants, because group one participants were compared to another player in their dossier. Providing another way of approaching the game may have caused group one participant to rethink their gameplay strategy.

Play Analytic Effects on Player Memory

I hypothesized that participants shown their data in-between games would have an easier time remembering what happened in the two games they played. This hypothesis turned out to be false. The participants who were shown the player dossier in-between games did no better, sometimes worse, than the participants who were not shown the player dossier.

A series of seven questions were used to test how well each participant remembered results from their gameplay. The first six questions were asked of every participant, whether they saw the dossier between games or not. Each participant was asked to compare their first and second play-throughs of Dihar. Participants had to decide, based on six different concepts found in Dihar, if they had performed better or worse in game one or two. The six questions asked were:

- 1) Did you score more prestige points in your first or second game?
- 2) Did you win more cases in your first or second game?
- 3) Did you have more clients in your first or second game?
- 4) Did you use more amendments in your first or second game?
- 5) Did you buy more upgrades in your first or second game?
- 6) Did you turn away more clients in your first or second game?

The participant's answers were cross-referencing with their gameplay data to determine if they had performed better/worse in one game compared to another. If players had equal performances in game one and two (e.g. if a player earned equal

amounts of prestige during both games) then those answers were removed as null answers. Of the question answers that were left, group three participants had the most correct answers, answering 73% of the questions correctly (65 correct out of 85). Group one was close behind with 71% correct answers (67 out of 94) and group 2 was further behind with 62.5% correct answers (70 out of 112). Even though group one and two were shown their player dossiers in-between games this did not seem to have had any affect how whether the participants were able to recall how they performed from one game to the next.

The final, seventh, question regarding a participant's memory was asked only of group one. Participants in group one were given a player dossier that compared some of their gameplay data with the data from a fake player. Group one participants were asked to recall if they had performed better than the fake player based on what they saw as part of the dossier. Most of the group one participants did perform better than the fake player based on the amount of prestige points earned. Only three participants did worse than the fake player in both games and while most of the participants knew they did better than the fake player most of them did not answer the question correctly. Group one participants were given the following choices when asked how they performance compared to the fake player: 'In the first game', 'In the second game', 'Both games', and 'Neither game'. Only 35% of the participants in group one (7 of 20) answered the question correctly. Enough players thought they had done better than the fake player but they didn't know which games they had done better. While it seems play analytic systems can be useful for reflecting on games, such as making players re-think their gameplay strategy, they do not help players remember their gameplay.

Perceptions of Play Analytics

Besides multiple choice questions, participants were also asked to provide full text answers regarding their perception of the player dossier system. Participants could write anything they wished as a response to the following questions:

1) When you viewed your profile in between games did any part of the dossier stand out? Why or why not?

2) Do you believe comparing your gameplay to another player in your dossier was helpful? Why or why not?

Both group one and group two were asked the first question, and only group one was asked the second question (since group one was the only group shown the player comparison part of the dossier). Also, the term profile was used in the study to refer to the player dossier because profile is a more common term. ‘Profile’ has been replaced with ‘dossier’ for consistency with the rest of the dissertation.

For the first question, half of the group one participants mentioned they noticed something in the dossier and referenced the fact that they needed to improve their skills. For example, one participant wrote

“The comparison between my dossier and the other [player] stood out because it seemed to show how badly I had done and so I knew second time around I had to do better.” (user ID: 1008)

Other participants mentioned how they did better than the other player or wrote about other differences between their two games.

“Upgrades for the waiting [area] stood out. I had way more [than the other player].” (user ID: 1073)

The other half of the group one participants said nothing stood out from the dossier and felt the dossier was not worth paying attention to. Answers like “Not really” or “I did not pay attention” were common among this subset of participants. A few reasons players gave for not caring about the dossier included: the dossier didn’t represent the player accurately, the participant already had a strategy they were relying on, and the participant did not want to compare themselves against the other player.

“Nothing stood out, I wasn't trying to compete with that player” (user ID: 1016)

The difference between the answers group one and two gave is how the ‘Yes the dossier stood out’ participants used the dossier for comparing gameplay. Group one participants tended to compare themselves to the other player (the fake player data provided), saying they needed to improve, or they had done better than the other player. Group two participants, who did not have access to another player’s data, mention they used the dossier to compare their own two games together; not to say they needed to improve but that they played differently in one game compared to the other. This points to the effects the player comparison had on the participants. When players are compared in a dossier (i.e. a player dossier) this alters how players perceived their own gameplay. When the study participants were given access to another player’s data the participants related the data to their own gameplay. Group two participants had no other player to compare with and they began comparing their own data, instead. For instance,

“Yes. I did substantially better during the first game. I won more cases and I lost fewer cases.” (user ID: 1032)

Additionally, almost all of the group two participants that said ‘nothing in the dossier stood out’ did not stand out gave no further explanation (only one participant mentioned the dossier didn’t show enough data). Conversely, group one responses, from the group one participants who said nothing about the dossier stood out, at least mentioned the comparison between players, even though those group one participants did not use the comparison. Having player comparison may be a better way for forcing players to acknowledge other methods for playing when compared to only showing players their own data.

The second question, posed only to group one the participants, asked participants if comparing their data with another player on their dossier was helpful. Eight out of 20

group one participants said the comparison was helpful and gave similar answers they gave to the first question. Some players mentioned they used the comparison as encouragement to do better, other players said they used the comparison to know they were performing well. One player specifically used the comparison to come up with a new gameplay strategy,

“Yes, in that I knew to buy more judges and upgrades sooner.” (user ID: 1084)

Of the group one participants that did not find the dossier helpful, half didn’t pay attention to it.

“Not really. I didn't really care that much about his [data].” (user ID: 1009)

The other half either mentioned they were not competing (and therefore the comparison didn’t matter) or they needed more data about the other player for the comparison to be helpful. These participants may have rejected the dossier because they perceived a bias in the system (comparing clients and upgrades). The bias did not match their way of playing, so the participants rejected the system. Overall, group one participants mentioned the player comparison feature even when they did not use the feature. Group two participants that rejected the dossier did so completely without acknowledging anything on the dossier, acting differently compared to group one.

Discussion

One purpose of conducting a remote study for this dissertation was to test the effects of a player dossier system, particularly how the system affects a player’s gameplay behavior and their perception of their gameplay. In order to explore if any effects existed I split my study participant population into three groups and gave them each a different player dossier system: group one saw both their own gameplay data and was compared to another “fake” player, group two saw only their own gameplay data and group three was not given a dossier in-between playing two games of ‘Do I Have A Right?’. I reiterate that the survey answer responses and gameplay variables collected

from the study participants were rarely found to be significant between groups. Even though group one and two were given a dossier system there were very little statistically significant effects compared to group three who never saw a dossier system containing gameplay data. I can therefore not definitively conclude that player dossiers, a form of play analytics, affect a player's gameplay behavior. However, the post-survey questions that solicited full text answers from the study participants do seem to show that, depending on what features are provided by a dossier system, a player's perception of their gameplay can be altered by the system. The evidence supporting this argument being many of the group one participants acknowledged the player comparison feature in the dossier system they were provided, even when they found the dossier unhelpful, and used the comparison to reflect on their gameplay strategies. By comparison, group two participants who did not find the dossier helpful provided no explanation why, completely ignoring the dossier. Also, other group two participants, those who acknowledge the helpfulness of the dossier, did not seem to change their behavior from game one to game two, at least when compared to group one. The player comparison feature seemed to have at least a small effect on how group one participants played their second game and how they perceived their gameplay in general.

Group one participants ended up performing better, or the same, than the participants in group two and three in their second games. They played longer than the other groups in the second game even though they were the quickest to finish the required four days. Also, group one participants were more likely to accept clients even when those clients should have been turned away. This may have been a result of the player comparison feature that visually compared clients and upgrades between players. As a result of accepting more clients group one participants also earned more prestige and lost less prestige during their second game while the other groups either did worse or stayed relatively static. Group one participants also bought more upgrades and changed their upgrade strategy more so than the other groups, which may have been, again, the result of comparing upgrades between players using the dossier.

The dossier system given to group one and two had a timeline visualization displaying data regarding a player's clients they helped in the game and the upgrades they bought over the course of the first four days of gameplay. The comparison portion of the dossier was intentionally designed to use graphics from the game (i.e client icons and icons representing the in-game upgrades) because the graphics/icons used related directly to what the participants experience in the game. A participant can see the items they bought and, if they were compared to another player, the items another player bought for comparison. The comparing feature was presented quite differently compared to the upper portion of the dossier. The upper portion displayed similar client and upgrade information just like the comparing feature. The number of cases won or lost was shown and a few bars were filled according given how many items the participant bought. However, when reviewing the full text answer questions many of the group one participants mentioned something about the dossier's player comparison and that they needed to improve their gameplay. I argue that the graphic comparison in the dossier made it easier for group one participants to understand the strategy used by the other player, more so than they would have had the information been presented in text/numeric form (i.e. similar to the upper portion of the dossier). Group one participants excelled at bringing in more clients and changing their upgrade strategy than either of the other two groups. As a result of understanding the other player's client and upgrade strategy, group one participants did better both by bringing in more clients and by altering their upgrade strategy.

Further research will have to be done in order to test whether comparing players using graphical depiction of data is better than text/numerical depictions. Perhaps a study can be conducted that provides different versions of a player comparison to a wider pool of subjects. Both graphical and numeric version would need to be tested. Staging an in-person user tests may also be beneficial in order to determine what players are thinking as they use a dossier system in-between gameplay. That may give a better indication of what players expect from a dossier system that compares player data.

The answers collected from the study's post-survey also provided a number of interesting play analytic related insights. For one thing, all of the study participants were willing to release their data. Most play analytic system release player data publically and, for this pool of participants, that does not seem to be a problem, although, the participants were hesitant when asked if their data could be connected to their real name. While most play analytic systems do not connect a player's real name to their data, Blizzard Ent. tried to enact a policy on Battle.net that would have link a player's real name to their Battle.net gamertag (Fletcher, 2010). Many Battle.net players protested the policy and Blizzard ultimately backed down, so the hesitation of this study's participants to reveal their real names, in connection with their data, seems to be shared among other play analytic player populations.

In regards to creating/sharing game content, participants said they were more likely to spent more time creating their character avatars and creating game content using in-game editors than capturing screenshots or sharing high scores. This may infer that concrete content (characters, items, etc.) related to a player's gameplay is more important to players than abstract concepts like high scores or screenshots. While many play analytic systems rely on presenting accumulated stats like scores or frequency of events, this approach may not be the best of presenting player data if play analytic systems are meant to keep players engaged with a game for longer. Systems that visualize game content, such as those that present player load-outs or content created using in-game editors, seem to be more likely to touch upon the aspects of gameplay players find the most personal. Furthermore, research also needs to be done as to how often players use APIs or MOD a game. Sharing and creating game content are related to APIs (which grants access to data) and MODs (which allows players to alter a game's content) so understanding why players use APIs or create MODs may provide further understanding of how a play analytic system can be built to foster sharing and creative behavior.

Realizing that graphs were not as preferred by the participants compared to other information formats like maps or graphics was fairly surprising. Most of the play analytic system presented in this dissertation use some form of graphs to present data and many

make extensive use of text/numerical values. It seems, however, that maps and graphics are preferred more so than other information formats, at least while playing a game. Although, there has been a recent rise in the number of map-based play analytic systems available. Map-based play analytics systems, like Halo Atlas or the Batman Arkham City Map application, are being released regularly and many of the larger play analytics systems (Bungie.net, Elite, Social Club) use maps as a means of visualizing game data. Graphics are often used as well: Bungie.net uses icons of in-game weapons, Gotham City Imposters use icons for their match awards and the Armory use icons/3D models for all the items represented in the game. It may have just been a fluke or players may have had misunderstood what graphs meant, but perhaps play analytic systems should prototype more map and graphic-based designs in addition to building systems that use graphs or numerical representations to visualize data.

Spectating games also seemed to be a fairly regular routine for many of the participants studied. Many of them spent, at least, a few hours a week spectating games in person or online. Up until this point, many of the play analytic systems described have been presented as systems for players to use, i.e. those who actually play the game the system was built for. But perhaps play analytic system should also be built to accommodate spectators too. There are examples of play analytic systems that provide replay analysis (SC2Gears, SC2Replays.com) and Halo Atlas can update a player's location in a multiplayer match in real time. These systems could be modified to provide better experiences for spectators, rather than only focusing on players alone. Spectators, using a play analytic system, could be provided real-time analysis of the match as it unfolds, allowed to watch instant replays or be able to work as support for the other player. The idea of giving spectators and players access to different capabilities isn't too farfetched. Games such as MAG, Natural Selections and Savage are designed to provide some players with unique command abilities versus the more combative, soldier abilities other players receive (for instance in Savage one player on a team would be able to build defenses or other buildings while the rest of the team attacked the enemies forces as if the game was a standard first-person shooter). Play analytic systems could provide spectators a way of taking part in a game as much as they could allow them to analyze the game

while they spectate. As we move to the conclusion of this dissertation next, I spend some time in the concluding chapter discussing topics such as spectating games and how play analytic systems can be designed to accommodate for a more diverse set of player experiences.

CHAPTER 7

CONCLUSION

On March 9th, 2012 the play analytic system built for the game Brink was shutdown. Unlike the similar shut down situation Halo Wars went through, no notification was given before the Brink system was taken down. There were no news stories about the shutdown and very little was said on the Brink forums (in fact the forum post announcing the system shutdown was locked and restricted players from commenting on the post). The reason the system was shutdown may have had to do with the limited number of players still playing Brink at the time of the shutdown, which was the same reason given for the Halo Wars shutdown (before the shutdown notices was rescinded). A dwindling player population may have also been the reason why there were no ‘high profile’ objections to the shutdown from Brink’s player community (unlike the same situation with the Halo Wars community). Brink was a new game under a new brand, unlike Halo Wars. Brink players may not have been as attached to the game, and the game’s world, enough to fight for the play analytic system to stay operational. Lack of players or support for the game, in addition to no warning of the impending shutdown, may have made it easier to remove support of the play analytic system without much outcry.

Brink’s play analytic system was available for nine months. When compared to other play analytic systems like Bungie.net, which was active for eight years, nine months is a short amount of time. Play analytics is a volatile and experimental domain. Other systems covered in this dissertation were non-operational when I began my work and just as many new systems were created during my work too. Developers are still trying to determine how these systems fit into a player’s experience along with a number of other factors surrounding gaming including: downloadable content (DLC), free-to-play business models, micro-transactions, social networking, and trans-media tie-ins. Games are no longer contained experiences that come on a cartridge or a disk and developers

must continually support the surrounding factors like play analytic systems. Many of these factors the game industry is struggling to monetize and improve revolve around games shifting from being goods to being services.

‘Gaming as a service’ is a common phrase used throughout the game industry today. The phrase symbolizes games ascension from being perceived as stand-alone software products to being systems game developers continue to support after the initial development phase is completed. Developers continue to develop content after a game is released and often charge extra money for continued access to new maps or other game content. Free-to-play games, which waive the initial cost of buying a game and instead offer to sell players smaller virtual items or features in the game, must maintain purchasing systems to keep track of players buying game content and effectively turns game development into a type of e-commerce. There are also digital distribution services like Amazon Digital Download that offer players the ability to buy and download games instantly. These are all services designed to give players better access to games and gameplay content.

Play analytics is another aspect of the ‘gaming as a service’ mentality affecting game development and is part of a growing trend to extend the experience of playing games beyond a game’s environment. By acting as a bridge between gameplay, game content and a player’s online identity, play analytics is one part of a process that takes the experience of playing games out of a game’s environment. When play analytics, and other communication features like forums or social networks, are combined with other gaming services, like digital distribution, we get hybrid service systems. Steam, for instance, allows players to purchase games digitally, form player groups and communicate through forums, in addition to tracking gameplay statistics and achievements. One might say that Steam and other play analytic systems are almost trying to move beyond just a service and instead provide an entire external experience for players. In this respect, the experience of gaming no longer refers to gameplay alone and now encompasses the act of playing, a game’s extended player community and the data/information/knowledge that exists around the game.

In their book “Authenticity: What Consumers Really Want”, authors Gilmore and Pine argue that there is a progression of stages a business can operate within based on what that business sells: commodities, goods, services, experiences (Gilmore and Pine, 2007). Gilmore and Pine state that as businesses progress through the stages they add the ability to customize and/or add authenticity in order to reach the next stage, such as customizing a commodity through diverse branding turns the commodity into a good. Adding customization is the process of making a stage more personal to the customer or more modifiable. For example, coffee (a commodity) is customized first into packaged good customers can choose from. Next, there are many businesses that brew coffee as a service allowing customers to choose where to go and how their coffee is prepared; a further example of customization. Finally, a company like Starbucks takes the first three stages and wraps them into a ‘third place’ experience (a place other than home or work) where customers not only are able to purchase coffee but are able to have a meaningful experience too. Starbucks customizes the service of brewing coffee by giving patrons a personalized space to have conversations and have personalized experiences while the business is selling the other stages (commodities, goods and services) as part of that experience. Authenticity, as Gilmore and Pine argue, factors into the four stages, much like customization, by referring to the way in which the stages are perceived by a customer: are the commodities seen as “natural” (was the coffee grown somewhere exotic?), are the goods “original” (does a coffee brand’s packaging come across as modern or helpful?), are the services “exceptional” (can a worker brewing a cup of coffee quickly and take care when making adjustments?) and are the experiences “referential” (does the Starbucks lounge remind customers of their living room or another comfortable place?). In a similar way to the functional/phenomenological framework I have used throughout this dissertation, progressing through the four stages requires businesses to customize what they are selling (a functional process) and be mindful of the authenticity of their products or services (a phenomenological process).

‘Gaming as a service’ is the latest step for the game industry along the progression model laid out by Gilmore and Pine. Digital games, in the past, have been

considered goods because they are built from code, art and other labor “commodities”. They are sold as products people can buy at stores. As the online connectivity of digital platforms increased (e.g PCs, gaming consoles, mobile devices, etc.), game developers and publishers have moved beyond just selling games to players. They now produce and provide direct services that modify and enhance the games being sold. Developers have built digital distribution services to sell and manage games, downloadable content systems to sell smaller pieces of game content, official communication platforms for players to converse and systems for tracking player’s gameplay (i.e player statistics). Players have been given greater customization through these services by letting players choose what to buy and offering them ways to connect with and compare themselves to other players. However, I do not regard play analytics as being another service attached to the ‘gaming as a service’ title, nor should we regard the rest of the services I mentioned as exist solely as services any longer. Play analytic systems like Bungie.net, Elite, Social Club and Steam are no longer merely providing services but are attempting to create whole experiences for players that exist outside of the game. These systems are highly customizable to players and combine many services (e.g. communication systems and digital distribution) together in one place. Play analytic systems also create “authentic” experiences by referring to the each player’s gameplay. Players are provided with meaningful accounts of their gameplay and a way to monitor their friends, or competitors, gameplay using play analytic systems. ‘Gaming as a service’ is now ‘gaming as a trans-experience’, one that acknowledges multiple types of experiences players can have regarding their favorite games, without having to play their favorite games in real-time. Players not only have the experience of gameplay (while they actually play the game) but also external experiences, which regularly contain play analytic systems. Although, what does play analytics offer as an external experiences for players? What are the categories or properties that define play analytics and how might these properties change over time as the domain of play analytics grows?

What Is Play Analytics?

A large part of this dissertation explores the state of the art of play analytic systems and has allowed me to define the properties play analytic systems offer players

as external experiences, separate (but connected) from their gameplay experience. Through the extensive content analysis performed in my work, I have learned that play analytics primarily offer players the data analysis methods to monitor, reminisce about and create data. Players monitor each other's data through leaderboards, activity feeds and groups. Recorded achievements and match stats allow players to reminisce about their past gameplay. Content generators allow players to create game content and APIs allow players to create play analytic systems for analyzing collected gameplay data. Each of these examples point to play analytic systems acting as places where players review their statistics, swap data, search for content and are ranked according to their gameplay.

In addition to learning which methods of data analysis are afforded by play analytic systems, I also found current play analytic systems revolve around increasing the level of competition in games, the accessibility of games and the ability for players to share content regarding their gameplay experiences. I arrived at these three principles after reviewing the systems I covered as part of this dissertation's content analysis and the results I collected from a small user study involving a play analytic system I created. Conflict, for example, appeared both in the play analytic systems reviewed and was referred to during the user study. Many of play analytic systems accompany competitive games in the FPS, RTS and RPG genres (such as Battlelog, Halo Wars and the WoW Armory) and offer the ability to compare or rank players against each other. Participants in the play analytic user study presented in chapter six also made mention of competition as being a driving factor in their desire to perform better in the game, after having seen their data compared to another player's data. Competition was also acknowledged by some participants who openly disregarded the desire to compete with other players because they felt playing the game their way was satisfactory.

The other two principles related to play analytics, accessibility and sharing, were also found in the user study results in addition to being common themes found amongst the play analytic systems covered. Participants in the user study who were shown another player's data were able to reassess the strategies they used during their gameplay based on the other player's data. Having another player's gameplay data allowed participants to

understand how they may approach playing the game differently. As part of the content analysis, accessibility is related to play analytic categories like content databases and maps; both provide searchable resources regarding game content and play gameplay data, acting like reference guilds or user manuals for players. The final property of sharing also relates to accessibility. Sharing another player's data made the game more accessible to participants during the user study and provided them with a way to reference how another player approached playing the game. I also found that sharing player data may have increased the participant's gameplay performance, allowing them to earn more points in the game studied, and seems to have made them more competitive. Although, players sharing data does not necessarily have to revolve around competition or accessibility in a play analytic system. Sending a screenshot to a friend in Modern Warfare 3's play analytic system, i.e. Elite, is just as easy as it is to compare their player stats. Sharing is about communication in addition to being a component of creating conflict and increasing accessibility.

The three common principles of play analytic systems – competition, accessibility and sharing – can be argued to mimic the design philosophies of many games developed today. Many designers and theorist reference the notion of competition, or artificial conflict, when defining games (Salen and Zimmerman, 2004). Furthermore, the accessibility of games has increased over the years for any number of reasons including: the desire of developers to sell more games to wider audiences, the experience players have with defined game conventions (referring to both control schemes and game mechanics), the increase attention paid to designing the user's experience happening in many digital-based professions, etc. Sharing, too, is a defining principle behind design movements like Web 2.0 (O'Reilly, 2005) and the abstract notion of sharing has only increased as more of the world is connected online. Yet, even while offering the means for players to compete, allowing players to share data and generally making games more accessible to players, play analytics rarely offers the one thing its name implies; the ability to analytically analyze game-related data.

What Play Analytics Can Be

Play analytic systems routinely offer players the ability to analyze their data. However, the analysis methods provided tend to allow players to monitor, reminisce or create their data. Players are not given the means to analytically analyze or reflect on their gameplay data or the content data found in games. Many current play analytic systems actually lie somewhere between the two analytic extremes discussed in the first chapter of this dissertation. Play analytics lies somewhere between normative and non-normative analytics. Below I provide reasons why play analytic systems are stuck somewhere between the two analytic extremes. I speculate as to how play analytics may move towards these two extremes simultaneously in the future and follow in the footsteps of other domains, such as information visualization, that have already been operating within the analytic extremes.

Play Analytics as Normative Analytics

Play analytic systems rarely act as typical normative analytic systems because players are restricted from asking questions and forming hypothesis when using play analytic systems. Play analytic categories like leaderboards and player stats tend to provide data in such a way that makes it hard to explore. It is hard to ask interesting analytic questions regarding a list of scores on a leaderboard or the few aggregated values from a player's stats. Other categories like global stats and match stats too may only provide aggregated values that hide the underlying data collected from larger groups of players. Relying on aggregated data takes away a player's ability to analyze gameplay collected from other players and use that analysis to reflect on their own gameplay. For instance, this dissertation's user study found that having the ability to compare a player's performance over time, not as aggregated values but as temporal events, did increase the likelihood of players mentioning they altered their play strategy based on the other player's data. However, many play analytic systems do not offer gameplay data historically and other systems offering game content are mainly used as reference systems (such as content databases), not analytic systems.

The inability to ask questions regarding a dataset means many of the analytic tasks common to the analytical data analysis method are not allowed in play analytics. For example, work done by Amar et al. found ten different analytic questions user typically ask when conducting analytical data analysis and the researchers related those questions to the following ten tasks (2005):

Table 7.1: Ten common analytic tasks people wish to perform when analyzing a data set.

<i>Retrieve Value</i>	Given a set of specific cases, find attributes of those cases.
<i>Filter</i>	Given some concrete conditions on attribute values, find data cases satisfying those conditions.
<i>Compute Derived Value</i>	Given a set of data cases, compute an aggregate numeric representation of those data cases.
<i>Find Extremum</i>	Find data cases possessing an extreme value of an attribute over its range within the data set.
<i>Sort</i>	Given a set of data cases, rank them according to some ordinal metric.
<i>Determine Range</i>	Given a set of data cases and an attribute of interest, find the span of values within the set.
<i>Characterize Distribution</i>	Given a set of data cases and a quantitative attribute of interest, characterize the distribution of that attribute's values over the set.
<i>Find Anomalies</i>	Identify any anomalies within a given set of data cases with respect to a given relationship or expectation, e.g. statistical outliers.
<i>Cluster</i>	Given a set of data cases, find clusters of similar attribute values.
<i>Correlate</i>	Given a set of data cases and two attributes, determine useful relationships between the values of those attributes.

Some of the analytic tasks found in Amar et al.'s research do appear in play analytic systems. Players can search for other players on a leaderboard or find an item on a content database. Players can filter usage stats or filter a leaderboard based on a set of criteria. Other tasks exist in play analytic system but are controlled and limited by the features provided in a system. Players can sort leaderboards but the leaderboards only contain the values designated by a game's developers or the creator of the leaderboard. Many of the variations in the stat categories compute derived values by aggregating player data and content generators derive content ratings by collect player votes; both are automated systems, however, where the system's creator defines what data is derived, not

the players using the system. Finally, some tasks almost never appear in play analytic systems. The ability to cluster, correlate or find anomalies are not tasks players are afforded given the currently available play analytic systems. There are cases where cluster or correlation may be used as part of a system but the features that make use of these tasks are not exposed to players. For example, Splinter Cell: Conviction's play analytic system has a player matching feature that analyzes a player's data before matching the player to opponents or potential co-op partners, but a player has no control over how the feature analyzes their data. Overall, some of the analytic tasks listed by Amar et al. are provided to play analytic users but only in a limited capacity, while other tasks are not offered at all.

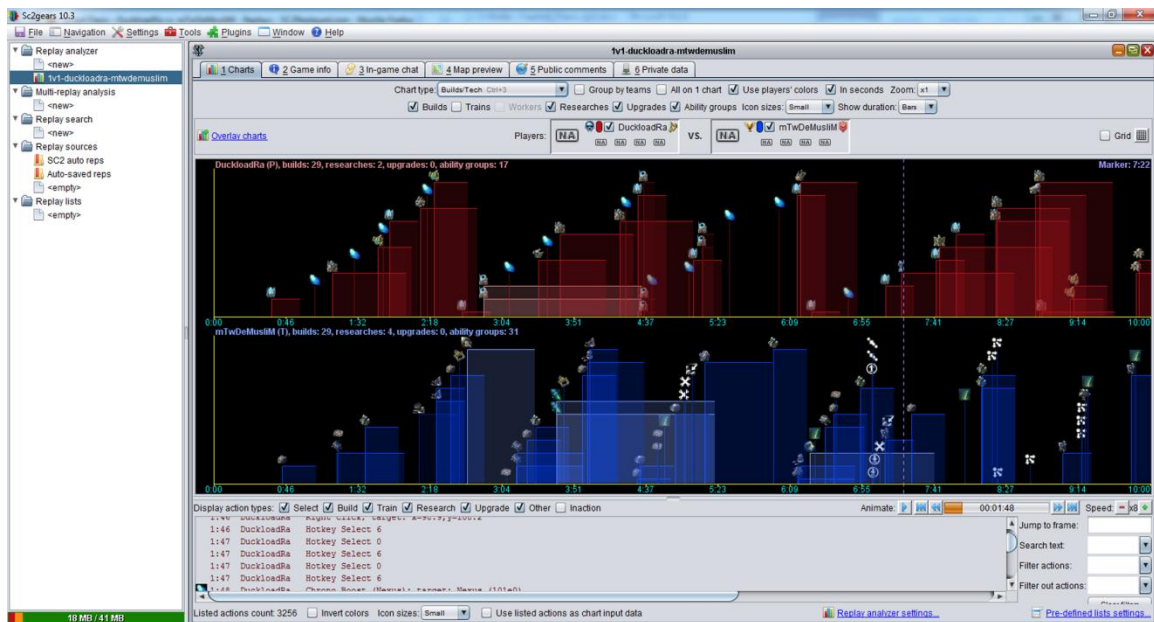


Figure 7.1: SC2Gears is one of the few examples of a normative play analytic systems.

Considering that play analytic systems provide players with a limited selection of the analytic tasks found in Amar's et al. research, we can speculate that future play analytic system can, and most likely will, provide more normative analytic approaches to analyzing game-related data. Offering players systems to analyze their performance and become better competitors is one way play analytics will likely continue towards the normative analytic extreme. I have argued throughout this dissertation that SC2Gears is one of the best examples of a play analytic system that helps Starcraft 2 players analyze their performance and fits into the definition of normative analytics (Figure 7.1).

SC2Gears affords many of the analytic tasks used in normative analytics: correlate

matches, find anomalies, determine range, derive values, etc. Players can analyze multiple matches at once in order to find pattern and analyze their gameplay compared against their opponent. SC2Gears can provide such analysis because it has access to Starcraft 2 replay files which contain a complete recording of the events that took place during a Starcraft 2 match. Replays are one of the few types of gameplay files that record historic data, data captured over time.. Systems like Bungie.net, Elite and Gotham City Imposters, all of which show kills and deaths over time, do not record and offer as much historic data at the same level available in replay files like the ones used for Starcraft 2. Although, if more play analytic systems were to offer detailed historic event data, similar to what replays offer, more systems like SC2Gears would likely be built by players wishing to analyze their gameplay.

As play analytic systems head towards the normative side of the analytics spectrum more of these systems may also appear as part of the real-time game experience. A few systems studied as part of the content analysis already exist as play analytic systems fused with a game environment. Autolog and RiderNet are friend leaderboard and recommendation systems built to coincide with the selection menus in their respective games, Need for Speed: Hot Pursuit and SSX. They work in a very similar way as other friend leaderboard system works except Autolog and RiderNet are accessed between gameplay sessions within the game environment. Other games with play analytics systems also make comparable stat systems available to player through various in-game menu systems, including: Modern Warfare 3, Gears of War 3 and Red Dead Redemption. External play analytic systems and their internal counterparts will continue to merge together following a similar merging precedence between play analytics and normative analytics. Once players are able to analytically analyze their gameplay externally, away from the game environment, it will only be a matter of time before players are given the ability to analytically analyze their gameplay in real time.

Play Analytics as Non-normative Analytics

Play analytics has not yet shifted far enough towards the opposite end of the analytic spectrum, away from normative analytics, towards embracing non-normative

analytics. Even though I argued in chapter one that the common interpretation of analytics, and later the common interpretation of visualization in chapter 3, follows a normative perspective (of utility and usefulness) there are examples of visualization domains operating with a non-normative mindset. The artistic, casual, critical and playful visualization domains discussed in chapter three all have examples of non-normative visualization tools and projects. Artistic work like Artifacts of the Presence Era (Veigas et al., 2004), playful systems like Fizz (bloom, 2011) and casual visualizations like Social Collider (Schmidt and Pohflepp, 2009) show that analyzing data can be provocative, frivolous and ambiguous. These systems may exist “for their own sake”, seem broken or not useful but still provide different data analysis perspectives and how data can be experienced. Play analytics too can take a similar route and expand into non-normative analytics, especially when we accept that play analytics can be qualitative and/or reflective.

Analyzing qualitative data already exists as part of some play analytic systems and often combines with the creation and sharing aspects of play analytics. Systems like Bungie.net and Elite have screenshot and video capturing features tied to their game’s environment. Players capture an image or video inside the game and it is made available through the play analytic system for players to share. Other players can rate the images or comment on videos but players can also keep the files private or only send them to friends. Someone could analyze these images analytically; they could analyze where images are often captured in the game, which images get the highest ratings or analyze the comments players leave on images. However, analyzing screenshots or videos can be a way players are given the ability to reminisce about their past gameplay. Players may capture exciting moments or hilarious game malfunctions causing unexpected results. Analyzing the data analytically would be counter-intuitive to the purpose of sharing gameplay videos and images. They exist ‘for their own sake’, for sharing interesting stories and collectively bonding over a shared appreciation of a game.

More play analytic systems will likely begin adding methods for players to create data “for its own sake”. One example that already exists, in addition to capturing

screenshots, is the story creation feature in the Sims Exchange. As part of the story creation feature, players capture screenshots while playing The Sims 3 and combine those shots with a fictional narrative they compose. These stories can be shared with other players in a similar way players are allowed to share images in Bungie.net or Elite. It is rare for play analytic systems to create narratives or stories around a player's gameplay but the Sims Exchange is one example where a player is given direct control over the creation of their own narratives. Players do not produce these narratives to analytically investigating their gameplay but can produce them for the sake of creating and sharing. In a way, the Sims Exchange story creation feature mimics the lifelogging systems I mentioned in chapter three. Projects like SenseCam Visual Diary (Lee et al., 2008) and MyTinyBits (Gemmell et al., 2006) capture images automatically as users proceed with their daily activities and allow users to annotate the images. Those projects can provide users with a normative means of analyzing their data, using their lifelogging data to comprehend and gain insights into their lives. But those same systems can be used for reminiscing and reflection, or other non-normative pursuits. One can even conceive of a game automatically capturing images from a player's gameplay, in a similar way as other lifelogging projects, and combining the images with gameplay data to create a play-by-play rendition without having to record an entire replay of the match. Play analytic systems are moving beyond being services that display player data to being systems that allow players to create content and experiences using their data. More qualitative features like capturing screenshots and creating narratives based on gameplay data will undoubtedly provide players with the means to produce meaningful artifacts that give them personalized experiences when using play analytic systems.

The second way play analytics may move towards non-normative analytics is to give players a chance to reflect on their gameplay and game content data. Reflecting on data gives players the ability to question how data is captured and for what reasons. Spore Skeletons, for instance, is one play analytic system that questions how player generated content should be visualized (Figure 7.2). Rather than present the actual Spore creature models, Spore Skeletons only visualizes the skeletal structure of each model. Players never see a model's skeleton in the game but the game's engine needs the skeleton data in

order to render the model correctly during gameplay. The Spore Skeletons project forces players to reflect on the creation process and makes players aware of what data is made available to them and what data is hidden from view.

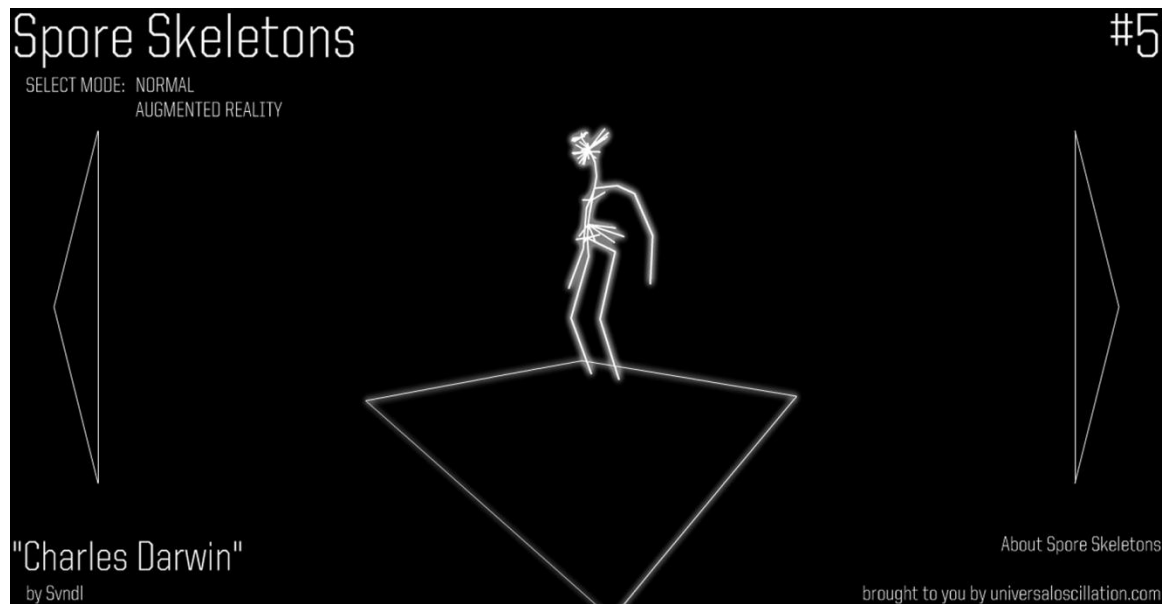


Figure 7.2: Spore Skeletons is one example of a non-normative play analytic systems.

Many play analytic systems do not give players direct access to their player data, or game content data, like the system built for Spore, which makes it possible for Spore Skeletons to function. This diminishes the capabilities of players to produce non-normative play analytic systems unless they work with a game developer directly or find a way to collect the data themselves. For example, some game developers create their own non-normative projects. Work done by Blackhurst, who produced the heatmaps found in the Just Cause 2 play analytic system, created a point cloud visualization using ‘death by impact’ locations data collected from Just Cause 2 players (2011). There were enough ‘impact’ points, over 11 million in total, that the visualization almost recreates the entire virtual island that exists in Just Cause 2 (Figure 7.3). The visualization is only presented as a video and focuses on the aesthetic value of the visualization rather than the analytical value the data may provide. However, Blackhurst is a developer that worked on Just Cause 2 and therefore had a much easier time gaining access to the data he needed to produce the point cloud visualization. If other, non-developer, players had the same level of access to game data as Blackhurst did they would likely produce similar non-normative work compared to Blackhurst’s visualization.

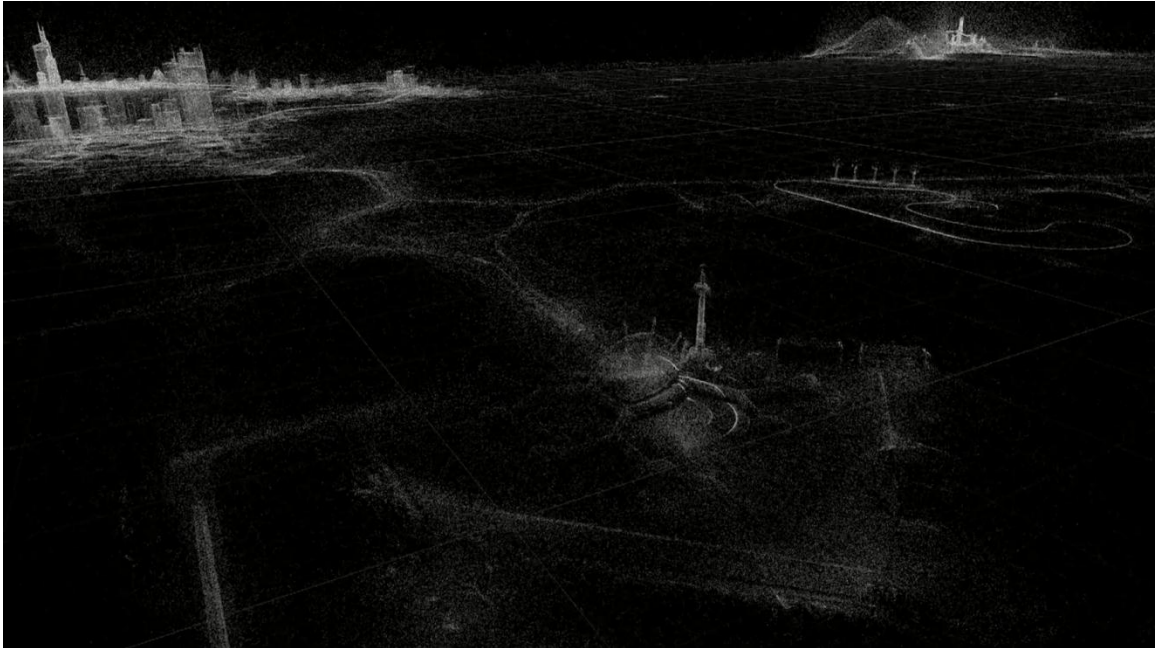


Figure 7.3: A point cloud visualization using player ‘death by impact’ locations found in Just Cause

2.

Another way players create non-normative play analytic examples, if they do not have direct access to any gameplay or content data, is by capturing and visualizing the data themselves. I already spoke about other play analytic examples like Darkfall Political maps and WOWhead that crowd-source data but these systems are examples of normative analytics. Data provided to Darkfall Political Maps, for instance, is used to give players a reliable map of the political alliances in the game world (making the game more competitive and accessible), not help players reflect on their data. Non-normative examples of projects that capture their own data include Huber’s Kingdom Hearts gameplay visualization (2010). Huber captured video footage from players who played Kingdom Hearts over a number of gameplay sessions. Video frames from that footage are used to visualize the gameplay sessions over time as single frames, taken at constant intervals, are positioned next to each other. Huber also visualizes the data in relation to how the player moves in the game environment, as seen in (Figure 7.4) where each row represents a frame from a different level in the game and each column represents the number of gameplay sessions over time. Additionally, one of my own projects follows a similar model of collecting gameplay footage as a method of obtaining data to visualize. As part of the project I too collected video of players playing JS Joust, a game played

within real space, but unlike Huber's project I combine the video footage recorded into long form exposure (Figure 7.5). The exposures allow me to visualize an entire game of JS Joust at once, revealing how players moved over time and when they interacted with one another. I could have built a system to record the spatial position of the motion controls each player carries during the game and used the data to create a similar visualization. Although, if I had used a standard data collection method some aspects of the visualization, such as a player's body posture or other none controller features of the game space, would have been erased. Even when players do not have direct access to gameplay data this has not stopped them from creating non-normative analytic systems. However, it is likely that if players were given greater access to gameplay data they would continue to develop non-normative systems just as much as they would produce normative systems.

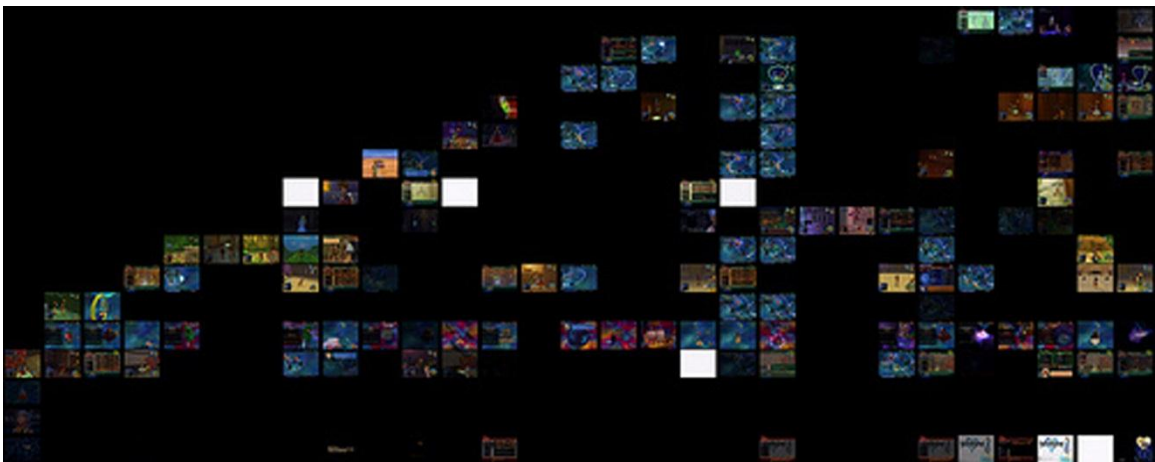


Figure 7.4: A visualization of video captured from multiple gameplay session of Kingdom Hearts. Each column represents a different play session and each row represents a particular world that exists in the game.

When players are not given access to gameplay/content data through traditional means (e.g. APIs or modding tools) they find ways to work around that limitation. Sometimes they use the data to build traditional, normative analytic tools like SC2Gears and other times non-normative like Spore Skeletons. Play analytics will grow in the coming years as part of the continued push by game developers to create external experiences, outside of gameplay, for players to engage with. Players will be given greater access to game-related data, either officially or unofficially, and they will use it to

create analytic systems that push towards the normative and non-normative extremes of the domain.



Figure 7.5: A long exposure shot taken of players playing the game JS Joust.

Dissertation Contributions

The contributions of this dissertation are four fold. First, projecting analytics simultaneously onto a normative and non-normative alignment is particularly relevant to the domain of games. There is a constant struggle between making games competitive, accessible and usable versus understanding that games are also about experimentation, escapism and fun. Players approach games for these competing reasons and in some cases one type of analytics, either normative or non-normative, is better suited for a player's needs or desires. When players want to analyze their gameplay in order to become better players they use a tool like SC2Gears, which follows a standard design for a normative analytic tool. When players want to share stories they use non-normative play analytic tools like the Sims Exchange to create narrated slideshows related to their gameplay. Presenting the dichotomy between normative and non-normative analytics has allowed me to conclude that play analytics exists somewhere in the middle of the two extremes and that future play analytic systems will likely simultaneously attempt to branch out towards those two extremes.

Second, the functional and phenomenological framework I use to interpret and critique play analytic systems offers a bridge between theory and practice. Bowker and Star's theories defining categorization systems are useful for interpreting play analytics as a kind of categorization system and my functional analysis of play analytics follow their three properties that define categorization systems: visibility, comparability and control. Visibility and comparability are properties often found in other definitions of visualization and analytics (Card et al., 1999; Thomas and Cook, 2005), however discussing who or what controls the data being analyzed is rarely discussed as an important aspect of analytic research. I argue that game developers have almost exclusive control over many play analytic systems and, at the very least, tend to control how data is collected and disseminated. This level of control may be stifling the creation of player-made play analytic systems. Furthermore, the phenomenological theories of Husserl, Heidegger and Gadamer are used as counter-points to Bowker and Star's three categorization properties in my theoretical framework. By using phenomenological theories regarding the experience we have with objects (data being the object in this case) I created three additional properties – exposure, interconnection and history – to aid me in my interpretation of play analytics as witnessed from the player's perspective. Many play analytic systems end up exposing a lot of player data and use that data to create interconnections between players, hoping those players will continue to share data and stay engaged by both the play analytic system and the game the system is built for. However, most play analytic system do not provide any meaningful historic data, opting instead to provide aggregated or up-to-date data related to a player's progression in a game. Lack of historical data reduces play analytic systems to be seen as tools for monitoring data instead of tools for analytically analyzing data over time (although, some tools like SC2Gears are exceptions). With more work, the framework used throughout this dissertation can be turned into a standardized approach for interpreting play analytic systems and may aid future play analytic system developers structure their play analytic designs.

Third, the content analysis presented in this dissertation is now a baseline for studying play analytics and represents a fairly complete state of the art look at the domain. There have always been colloquial terms for defining the systems and features that form the domain of play analytics: leaderboards, achievements, stat tracking, companion applications, etc. This dissertation formalizes these systems into a common domain complete with overarching categories and feature variations that exist in those categories. Those categories, and their variations, now act as design patterns that can be formally studied and explored within the domain of play analytics. Stat tracking now extends into multiple categories including player stats, match stats, group stats and global stats. Other system types thought to be separate from features like stat tracking, including player content generation and content databases, are also included as part of play analytics due to their connection to both player and game content data. I also found that many play analytic systems are built around three central themes: to provide players with the means of competing against one another, to allow players to share data with each other and to make a game more accessible by allowing player to reference game content outside a game's environment. These themes relate to the design principles behind game development today, seeing as many games promote competition, encourage sharing and have been made more accessible to wider audiences. However, these themes may also be the reason why play analytic systems exist somewhere in between the two analytic extremes of normative and non-normative analytics. Play analytic systems do not afford players the ability to analytically analyze their data (i.e. normative analytics) because the data provided to players often lack historical context. Instead, these systems are focused on sharing or comparing new data, leaving old data to lose its meaning. Nor do play analytics afford the ability to reflect and question how data is analyzed (a non-normative approach) because data is provided in order to make a game more accessible, not critically question how game data is collected or represented. I speculate that as the popularity of play analytics grows the domain will stretch towards both normative and non-normative extremes simultaneously as more traditional analytic tools are built to analyzes games (like SC2Gears) and players begin creating more critical or personalized systems (like Spore Skeletons).

Fourth, the user study conducted for this dissertation did not prove as conclusive as I had hoped but the results were not outright negative; even somewhat positive in some cases. The participants in group one, who were able to compare their data to another player after playing their first game in the user study, did perform better during their second game after they compared their data against another player's data. The other participants, both from the group that viewed only their own data and the group that was given no data at all, performed slightly worse during their second game. Group one participants also were found to be more likely to alter their gameplay strategy during their second game and some group one participants specifically referenced the player comparison as the reason for their gameplay alterations. Further studies will have to be conducted in order to determine if the perceived effects on group one are accurate because there were a number of reasons the data collected was found to be insignificant, including: the low population size sampled for each group, the narrow age range of the participants, the recruitment of participants from a single university class, the education level of the participants and the choice to studying a smaller browser-based game. Future studies will be able to eliminate these problems and below I discuss how other types of studies may be conducted to explore the effects play analytic has on a player's gameplay experience.

The Future of Play Analytic Research

Remote player studies, like the one conducted in chapter six, are just one type of study that can be performed in order to explore how play analytics affect a player's gameplay experience and how the future development of play analytic systems can progress. Ethnographic studies, for instance, would be much more preferred way to understand what players actually do while they use play analytic systems and how players feel about their experience. Play analytic systems are not necessarily cultural groups that a researcher can study but in-person lab observation, as well as observation of players in the environment they typically play games, could be beneficial for understanding how players interact with play analytic systems. It may also be to a play analytic researcher's benefit to monitor the groups that form within play analytic systems. Since play analytics relies heavily on competition and sharing, understanding how groups

compete and share data may be a good starting point for discovering how new play analytic features may accommodate larger groups of players. Besides ethnographic studies, applying traditional cognitive science, human-computer interaction and information visualization research methodologies would also benefit play analytic research. These fields already study how users memorize data, interact with interfaces and interpret data. A study similar to (Amer et al. 2005) mentioned above could be performed using play analytic systems, for example. Players could be asked what questions they have regarding their gameplay data and what they expect a play analytic system to provide. No doubt a similar set of tasks as the ones Amer et al. found would be highlighted, however there may be additional responses related to other data analysis approaches, like reminiscing or reflection, because not every player approaches their gameplay analytically. Applying traditional research methodologies like those found in information visualization research may be a good combination to expand upon play analytic research.

Other ways play analytic research may expand is by studying the development process behind play analytic systems. Play analytic systems are caught between being external to a game's environment but are meant to refer back to their connected games. This means play analytic development needs to draw on the design of their connected game and also rely on other design methodology related to the platforms these systems reside on. Considering many play analytic systems are made available over the web, design methodologies regarding web and server development are often used as part of the play analytic development process. Gaining a greater understanding of the capabilities of modern web browsers and servers would assist play analytic developers in building optimized and sustainable solutions for future play analytic systems. Future business models for play analytic development may also include payment packages for players that allow them to store their own personal data and it will be important for game developers to have proper methods for storing and archiving gameplay data over time. SC2Gears offers a payment package to players who wish to store large numbers of Starcraft 2 replay files online, for instance. Acknowledging how online development works and combining those practices with game development will likely be a requirement

for any game developer, or players, seeking to produce play analytic systems going forward.

Finally, there are additional systems and services that exist around game and relates to play analytics. These systems should be studied in combination with play analytics. First, many gaming groups have online guild, or group, websites where a guild will often keep track of their own gameplay data (sometimes manually collected) and serves as a place where the guild can converge and converse. Play analytic systems can serve as hubs where players congregate but researching player run guild websites may be beneficial to determine how the websites differ, and are similar to, play analytic system. Second, there are a number of video streaming websites and game streaming services that allow players to watch gameplay without having access to the game they are watching or a gaming platform to play the game. Services like Twitch.TV and Onlive let players stream their gameplay footage live and archive old streams for viewers to watch later. Some play analytic systems give users the ability to capture video (Bungie.net, Driver: San Francisco, Elite) but these video files are usually very short. Video streams on Twitch.TV can be hours long and may include voice overs of the players providing commentary of their gameplay. With the interest in e-sports (i.e. electronic sports that often refers to professional video game playing) growing in North America and Europe (along with other countries like Korea who have been interested in spectating e-sport gaming events for much longer) (Taylor, 2012) spectating video games is becoming as popular as playing them. Twitch.TV and Onlive offer a way for people to engage in spectating e-sports or merely regular gameplay made available by everyday players. Similarly, play analytic systems are also a way of spectating and analyzing gameplay. It is certainly possible that game spectating services like Twitch.TV will be combined with play analytic systems offering gameplay analysis in the future. Play analytic researchers should be monitoring and exploring how these related services are evolving over time in order to determine how these services may be combined with play analytic systems.

Play analytics is becoming a major part of the expanded experience surrounding games and continues to grow as a domain. Perhaps the most significant recent event

involving play analytic systems occurred during the fall of 2011 when Battlefield 3 and Modern Warfare 3 were both scheduled to be released. Each game was marketed as the best modern military shooter that year and one of the defining points being compared between the two games were their play analytic systems: Battlelog for Battlefield 3 and Elite for Modern Warfare 3. Battlelog was presented as a player hub for PC players, where the game could be launched and players could communicate with each other, while also pushing the fact that Battlelog was entirely free. Elite was presented as a useful tool for players to analyze and increase their gameplay performance, while the purchasable premium Elite membership gave players access to all future Modern Warfare 3 DLC.

Both Battlelog and Elite provide a glimpse as to where play analytics is heading. Play analytic system are now being designed to operate as compete external experience outside of the player's gameplay experience, much like how Battlelog functions. There is also an emphasis being placed on play analytic systems to provide players with better methods of analyzing their data, as Elite is said to provide. Additionally, paying for play analytic systems may become a way to ensure the systems remain functional for longer periods of time. With the influx of more attention and funding play analytics will grow much fast than it has over the last decade as more players are tracked and further access to game-related data is exposed. The research conducted as part of this dissertation is only a starting point to understanding what the play analytic domain can offer players. Further play analytic research is required as the domain evolves and this research will help designers, whether they are game developers or game players, develop future play analytic systems that allow players to play with data.

APPENDIX A

PLAY ANALYTIC SYSTEM INFORMATION

Below are tables containing the 81 systems I covered as part of the content analysis presented in chapter 5. The systems are ordered alphabetically and are split up into groups of 15 or less so they may fit on each page. Each system is also cross-referenced with the eight play analytic categories I found while conducting the content analysis. A blue box marks which categories a particular play analytic system had as part of the system's feature set. After the tables describing the systems are covered I list each system's online URL or, if the system is no longer active, I mention if the system has been shut down or inoperable.

Table A.1: Content Analysis Play Analytic Systems – 1 to 15

	Leaderboard	Player Stats	Group Stats	Match Stats	Global Stats	Maps	Content Generation	Content Database
Achieve 360 points								
AION								
Armor Games								
Army of Two: The 40th day								
Assassin's Creed 2 (Franchise)								
Bad Company								
Bad Company 2								
Batman Arkham City Map App								
Battle.net								
Beatles Rock Band								
Battlefield 3 – Battlelog								
Brink								
Bungie.net								
Burnout Paradise								
Call of Duty: Modern Warfare 3 - Elite								

Table A.2: Content Analysis Play Analytic Systems – 16 to 30

	Leaderboard	Player Stats	Group Stats	Match Stats	Global Stats	Maps	Content Generation	Content Database
Crimecraft								
Darkfall Political Maps								
Day Of Defeat								
Demigod								
Driver: SF								
EA Football Club								
EA sports hub								
EA sports Active								
End War								
Eve Online								
FIFA Earth								
Forza 4 – My Forza								
GamerDNA								
Gears of War 3 – War Journal								
Giant Bomb								

Table A.3: Content Analysis Play Analytic Systems – 31 to 45

	Leaderboard	Player Stats	Group Stats	Match Stats	Global Stats	Maps	Content Generation	Content Database
Gotham City Impostors								
Greenday Rock Band								
Guild Wars								
Half Life 2								
Halo Wars								
Halo Waypoint								
Halo Atlas								
Heroes of Newerth								
Infamous 2								
Just Cause 2								
Kongregate								
Little Big Planet Series								
League of Legends								
Lord of the Rings								
MAG								

Table A.4: Content Analysis Play Analytic Systems – 46 to 60

	Leaderboard	Player Stats	Group Stats	Match Stats	Global Stats	Maps	Content Generation	Content Database
Map WoW								
Maplestory								
Mod Nation Racer								
Mortal Kombat								
My Killzone								
Need For Speed: Hot Pursuit – Autolog								
NewGrounds								
Noby Noby Boy								
Orcs Must Die								
Planetside								
Play Station Trophies								
PSNProfiles.com								
Red Faction								
Resistance 3								
RiderNet								

Table A.5: Content Analysis Play Analytic Systems – 61 to 75

	Leaderboard	Player Stats	Group Stats	Match Stats	Global Stats	Maps	Content Generation	Content Database
Saints Row 3								
SC2Replayed								
Settlers 7								
Sims 3 Exchange								
Skate 3								
Social Club – LA Noire								
Social Club – Red Dead								
Splinter Cell: Conviction								
Sporepedia and API								
Spore Skeletons								
Starcraft 2– SC2Gears								
Steam								
Super Create Box								
Terraria Map								
The Godfather 2								

Table A.6: Content Analysis Play Analytic Systems – 76 to 81

	Leaderboard	Player Stats	Group Stats	Match Stats	Global Stats	Maps	Content Generation	Content Database
Torchview								
Warhammer 40k								
WoW Armory								
WoW Auction App								
WoW Head								
Xbox Live Website								

Content Analysis Play Analytic Systems – References

Achieve 360 points - <http://www.achieve360points.com/>

AION - <http://na.aiononline.com/>

Armor Games - <http://armorgames.com/>

Army of Two: The 40th day - <http://www.armyoftwo.com/home.action>

Assassin's Creed 2 (Franchise) - <http://assassinscreed.uk.ubi.com/assassins-creed-2/>

Bad Company – (No longer operational)

Bad Company 2 - <http://www.battlefieldbadcompany2.com/>

Batman Arkham City Map App - <http://itunes.apple.com/us/app/batman-arkham-city-official/id472027508?mt=8>

Battle.net - <http://us.battle.net/en/>

Beatles Rock Band - <http://thebeatlesrockband.com/>

Battlefield 3 – Battlelog - <http://battlelog.battlefield.com/>

Brink – (No longer operational)

Bungie.net – (No longer collecting data) - <http://www.bungie.net/>

Burnout Paradise - <http://www.criteriongames.com/burnout/paradise/>

Call of Duty: Modern Warfare 3: Elite - <http://www.callofduty.com/elite>

Crimecraft - <http://www.crimecraft.com/>

Darkfall Political Maps - <http://darkfallinfo.com/pmap/>

Day Of Defeat - <http://www.dayofdefeat.com/stats/>

Demigod - <http://www.demigodthegame.com/>

Driver: SF - <http://driver-thegame.ubi.com/driver-san-francisco/en-US/home/>

EA Football Club - <http://www.ea.com/uk/football/>

EA sports hub - <http://www.easports.com/>

EA Sports Active – (No longer operational)

End War – (No longer operational)

Eve Online - <https://gate.eveonline.com/>

FIFA Earth – (No longer operational)

Forza 4: My Forza - <http://forzamotorsport.net/en-us/Profile.aspx>

GamerDNA - <http://www.gamerdna.com/>

Gears of War 3: War Journal - <http://gearsofwar.xbox.com/en-US/warjournal>

Giant Bomb - <http://www.giantbomb.com/>

Gotham City Impostors - <http://www.gothamcityimpostors.com/>

Greenday Rock Band - <http://www.greendayrockband.com/>

Guild Wars - <http://guildwars.com/>

Half Life 2 - http://www.steampowered.com/status/ep2/ep2_stats.php

Halo Wars - <http://www.halowars.com/>

Halo Waypoint - <http://halo.xbox.com/en-us/?>

Halo Atlas - <http://halo.xbox.com/en-us/Atlas>

Heroes of Newerth - <http://www.heroesofnewerth.com/>

Infamous 2 - http://www.infamousthegame.com/en_US/infamous2.html

Just Cause 2 - <http://www.justcause.com/>

Kongregate - <http://www.kongregate.com/>

Little Big Planet Series - <http://www.littlebigplanet.com/en/>

League of Legends - <http://na.leagueoflegends.com/>

Lord of the Rings - <http://www.lotro.com/>

MAG - (No longer operational)

Map WoW - <http://mapwow.com/>

Maplestory - <http://maplestory.nexon.net/>

Mod Nation Racer - <http://community.modnation.com/en-us/>

Mortal Kombat - <http://www.themortalkombat.com/>

My Killzone - http://www.killzone.com/kz3/en_US/mykillzone/banner.html

Need For Speed: Hot Pursuit: Autolog - (Available in game) -
<http://www.needforspeed.com/autolog>

NewGrounds - <http://www.newgrounds.com/>

Noby Noby Boy - <http://apps.evilrobotstuff.com/nobynobyboy/>

Orcs Must Die - <http://www.robotentertainment.com/games/orcsmustdie/>

Planetside - (No longer collecting data) - <http://planetside.station.sony.com/>

Play Station Trophies - <http://www.ps3trophies.org/>

PSNProfiles.com - <http://psnprofiles.com/>

Red Faction - <http://www.redfaction.com/>

Resistance 3 - <http://resistance-game.com/>

SSX: RiderNet - (Only system available only in the game)

Saints Row 3 - <http://www.saintsrow.com/>

SC2Replayed - <http://www.sc2replayed.com/>

Settlers 7 - <http://thesettlers.us.ubi.com/the-settlers-7/my-realm/>

Sims 3 Exchange - <http://www.thesims3.com/exchange>

Skate 3 - <http://skate.ea.com/>

Social Club: LA Noire - <https://socialclub.rockstargames.com/games/lan/index-signin.html>

Social Club: Red Dead - <https://socialclub.rockstargames.com/games/rdr/index-signin.html>

Splinter Cell: Conviction - <http://g2w.ubi.com/splinter-cell-conviction/match/en/users>

Sporepedia and API - <http://www.spore.com/sporepedia>

Spore Skeletons - <http://www.universaloscillation.com/sporeskeletons/>

Starcraft 2: SC2Gears - <https://sites.google.com/site/sc2gears/>

Steam - <http://store.steampowered.com/>

Super Create Box - <http://supercratebox.com/>

Terraria Map - <http://www.terrariaonline.com/threads/terraria-map-viewer.4226/>

The Godfather 2 - (No longer operational)

Torchview - <http://torchview.codeplex.com/>

Warhammer 40k - <http://www.dawnofwar2.com/us/home>

WoW Armory - <http://us.battle.net/wow/en/>

WoW Auction App - <http://us.battle.net/wow/en/services/wow-remote/>

WoW Head - <http://www.wowhead.com/>

Xbox Live Website - <http://www.xbox.com/en-US/live/>

REFERENCES

- Aarseth, E. 1997. *Cybertext: Perspectives on Ergodic Literature*. JHU Press.
- Ackoff, R.L. 1989. From Data to Wisdom. *Applied Systems Analysis*. 16, p. 3–9.
- Activision Publishing Inc. 2011. Elite (Call of Duty: Modern Warfare 3). Accessed January 9, 2012. <http://www.callofduty.com/elite>
- Albert, J. 2007. Hitting in the Pinch. In: Albert, J. and Koning, R. H. (eds.) *Statistical Thinking in Sports*. CRC Press, p.111-134.
- Amar, R., Eagan, J., Stasko, J. 2005. *Low-level components of analytic activity in information visualization*. In Preceedings of Information Visualization 2005. p. 111-117.
- Amar, R. and Stasko, J. 2005. Knowledge Precepts for Design and Evaluation of Information Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4), p.432-442.
- Anderson, J. 2004. *Cognitive Psychology And Its Implications*. Macmillan.
- Baker, S. 2009. *The Numerati*. Houghton Mifflin Harcourt
- Beizer, B. 1990. *Software Testing Techniques*. University of Michigan Press.
- Belicza, A. 2010. Sc2gears. Accessed on Sept. 15, 2010, <http://sites.google.com/site/sc2gears/>.
- Bergeron, B.P. 2002. *Dark Ages II: when the digital data die*. Prentice Hall.
- Blackhurst, J. 2011. *Heatmaps, Point Clouds and Big Data in Processing*. Accessed May 25, 2012. <http://jimblackhurst.com/wp/2011/05/17/heatmaps-point-clouds-and-big-data-in-processing/>

- Blizzard Ent. 2010. World of Warcraft Armory. Retrieved April 24, 2011 from <http://us.battle.net/wow/en/>.
- bloom. 2011. Fizz. Accessed on April 15, 2012. <http://fizz.bloom.io/>
- Bogost, I. 2007. *Persuasive Games: the expressive power of videogames*. MIT Press.
- Bowker, G. and Star, S. 2000. *Sorting Things Out: classification and its consequences*. MIT Press.
- Brin, D. 1998. *The Transparent Society: will technology force us to choose between privacy and freedom?* Basic Books.
- Brown, S. and Vaughan, C. 2009. *Play: How It Shapes the Brain, Opens the Imagination, and Invigorates the Soul*. New York: Penguin Group.
- Bogost, I. and Montfort, N. 2009. Platform Studies. Accessed on April 12, 2012. <http://platformstudies.com/>
- Boyd, D. 2010. *Making Sense of Privacy and Publicity*. Accessed on February 15, 2011, <http://www.danah.org/papers/talks/2010/SXSW2010.html>.
- Budd, R., Thorp, R., Donohew, L. 1967. *Content analysis of communications*. Macmillian.
- Bungie LLC. 2004. Bungie Online. Accessed on December 1, 2010, <http://www.bungie.net/online/default.aspx>.
- Bungie LLC. 2010. Halo: Reach API. Accessed on February 15, 2011, <http://www.bungie.net/fanclub/statsapi/group/grouphome.aspx>.
- Butler, J. 2004. *Undoing Gender*. Psychology Press.
- Byrne, D., Lee, H., Jones, G., Smeaton, A. 2008. *Guidelines for the presentation and visualisation of lifelog content*. In proceedings of Irish Human Computer Interaction Conference 2008.

- Caillois, Roger. 2001. *Man, Play and Games*. Translated by Meyer Barash. Illinois: University of Illinois Press. [1961].
- Cadora, E. and Kurgan, L. 2006. *Architecture and Justice* (exhibition brochure). The Architectural League of New York. September 15-October 28, 2006. Accessed on April 15, 2012.
<http://www.spatialinformationdesignlab.org/publications.php?id=38>
- Card, S., Mackinlay, J. and Shneiderman B. 1999. *Readings in Information Visualization: Using Visualization to Think*. Morgan Kaufmann.
- CCP. 1997. Eve Online API. Accessed on February 15, 2011,
<http://www.eveonline.com/api/doc/>.
- Chen, M., Ebert, D., Hagen, H., Laramée, R.S., van Liere, R., Ma, K.-L., Ribarsky, W., Scheuermann, G., Silver, D. 2009. Data, Information, and Knowledge. In *IEEE Visualization, Computer Graphics and Applications*. 29(1), p.12-19.
- Cocopjojo (alias). 2010. Halo Waypoint Transition (Please Read). Halo Wars Forum. Accessed on April 12, 2012.
<http://www.halowars.com/forums/1/953396/ShowThread.aspx>
- Consalvo, M. 2007. *Cheating: Gaining Advantage in Videogames*. MIT Press.
- Consalvo, M. 2011. *Using your friends: Social mechanics in social games*. In Proceedings of Foundations of Digital Games 2011.
- Corner, J. 1999. The Agency of Mapping. In *Mappings*, ed. Cosgrove, D. p. 213-52.
- Crawford, C. 2003. *Chris Crawford on game design*. New Riders.
- Davenport, T., and Harris, J. 2007. *Competing on Analytics: The New Science of Winning*. Harvard Business School Press.
- Darkfall Info. 2006. Darkfall Information Database. Accessed on February 15, 2011,
<http://darkfallinfo.com/>.
- Data.gov, 2010. Accessed on January 2, 2011, <http://www.data.gov/>.

- Dawe, M., Graham, R., Schwab, B. 2011. AI Pr0n: Maximum Exposure of Your Debug Info! Presented at the Game Developers Conference 2011
- de Certeau, M. 1984. *The Practice of Everyday Life* (English translation). Trans. Steven Rendall. University of California Press.
- Derrida, J. 1998. *Archive Fever: A Freudian Impression*. University of Chicago Press.
- Desai, Rushi. *New Tools to Optimize App Authentication*. Facebook Blog. Accessed on October 20, 2011, <https://developers.facebook.com/blog/post/578/>
- Devlin, H. (2008). *What is Functional Magnetic Resonance Imaging (fMRI)?* Psych Central. Accessed on June 16, 2012. <http://psychcentral.com/lib/2007/what-is-functional-magnetic-resonance-imaging-fmri/>;
http://psychcentral.com/lib/img/fmri_scan.jpg
- Dewey, J. 1933. *How We Think: A restatement of the relation of reflective thinking to the educative process*. Boston: D.C. Heath.
- Ducheneaut, N., Yee, N., Nickell, E., and Moore, R.J. (2006). *Alone Together? Exploring the Social Dynamics of Massively Multiplayer Games*. In conference proceedings on human factors in computing systems CHI 2006, pp. 407-416. April 22-27, Montreal, PQ, Canada.
- Dourish, P. 2004. *Where The Action Is: The Foundations Of Embodied Interaction*. MIT Press.
- Edmunds, A. and Morris, A. 2000. The problem of information overload in business organizations: a review of the literature. *International Journal of Information Management*. 20, p.17-28.
- Electronic Arts. 2004. The Sims Exchange. Accessed on June 5, 2011. <http://www.thesims3.com/exchange>.
- Electronic Arts. 2010. Autolog. Accessed on December 1, 2010. <http://hotpursuit.needforspeed.com/game-info/autolog>.

- Eppler, M. and Mengis, J. 2004. The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*. 20, p.325–344.
- Erikson, E. 1968. *Identity, youth and crisis*. New York: Norton.
- Farmer, R. and Glass, B. 2010. *Building Web Reputation Systems*. O'Reilly Media Inc.
- Felton, N. 2005. *Annual Report (2005-2011)*. Accessed on Aprle 15, 2012.
<http://feltron.com/>
- Few, S. 2009. *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. Analytics Press.
- Fletcher, JC. 2010. *Blizzard's Battle.net director on Real ID, cross-game integration and BlizzCon*. Accessed on May 8, 2010.
<http://www.joystiq.com/2010/10/21/blizzards-battle-net-director-on-real-id-cross-game-integratio/>
- Floridi, L. 1999. *Philosophy and Computing: An Introduction*. Psychology Press.
- Fong, D. 2007. Raptr. <http://raptr.com/>
- Friendly, M. and Denis, D. 2009. *Milestones in the history of thematic cartography, statistical graphics, and data visualization*. Accessed on April 15, 2012.
<http://www.datavis.ca/milestones/>
- Fry, B. J. 2004. *Computational Information Design*. Doctoral Dissertation. Massachusetts Institute of Technology, Cambridge.
- Gadamer, H. G. 1994. *Truth and Method*. Trans. Joel Weinsheimer and Donald G. Marshall. New York: Continuum.
- Galloway, A. R. 2006. *Gaming: Essays On Algorithmic Culture (Electronic Mediations)*. Univ Of Minnesota Press.
- Galotti, K. 2010. *Cognitive Psychology: In and Out of the Laboratory*. Cengage Learning.

- GameTracker. 2005. GameTracker.com. Accessed on February 15, 2011 at <http://www.gametracker.com/>.
- Ganesh, A. 2010. The essential Chennai rail transport map – looking back. Accessed on January 2, 2011, <http://bitterscotch.wordpress.com/2010/05/07/the-essential-chennai-rail-transport-map-looking-back/>.
- Gee, J.P. 2007. *What video games have to teach us about learning and literacy*. Palgrave Macmillan, New York.
- Gemmell, J., Bell, G., Lueder, R. 2006. MyLifeBits: a personal database for everything. *Communications of the ACM - Personal information management*, 49(1). p. 88-95.
- Gerstmann J. and Davis, R. 2008. Giant Bomb. Accessed on June 5, 2011. <http://www.giantbomb.com>.
- Gilliom, J. 2001. *Overseers of the poor: surveillance, resistance, and the limits of privacy*. University of Chicago Press.
- Gilmore, J. and Pine, B. 2007. *Authenticity: what consumers really want*. Harvard Business School Press.
- Goodman, P. 1993. *Practical Implementation of Software Metrics*. McGraw-Hill.
- Guillaume and Joshua. 2006a. Wowhead. Accessed on February 12, 2011, <http://www.wowhead.com/>.
- Guillaume and Joshua. 2006b. Wowhead client. Accessed on February 12, 2011, <http://www.wowhead.com/client>.
- Hall, P. 2008. Critical Visualization. In *Design and the Elastic Mind*, by Aldersey-Williams, H., Hall, P., Sargent, T., and Antonelli, P., D.A.P./Distributed Art Publishers, Inc. p. 122-131.
- Halo Wars. 2009. Ensemble Studios (No Longer Operating). Microsoft Studios. Xbox 360.

- Halo Wars Sales Figures. 2012. VGChartz.com. Accessed on April 12, 2012.
<http://www.vgchartz.com/game/7715/halo-wars/>
- Hatton, N. and Smith, D. 1995. Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, 11(1), p.33-49.
- Heer, J. and Bostock, M. 2010. *Crowdsourcing graphical perception: using mechanical turk to assess visualization design*. Proceedings of the 28th international conference on Human factors in computing systems.
- Heeter, C., Lee, Y., Magerko, B. and Medler, B. 2011. Impacts of Forced Serious Game Play on Vulnerable Subgroups. *International Journal of Gaming and Computer-Mediated Simulations*. 3(3).
- Heidegger, M. 1962. *Being and Time*, Trans. by John Macquarrie and Edward Robinson. New York: Harper & Row.
- Helix Systems Inc. 2003. WoWecon. Accessed on February 15, 2011,
<http://www.wowecon.com/>.
- Herbrich, R. and Graepel, T. 2006. TrueSkill(TM): A Bayesian Skill Rating System, no. MSR-TR-2006-80. Accessed on April 12, 2012.
<http://research.microsoft.com/apps/pubs/default.aspx?id=74419>
- Holsti, O. 1969. *Content Analysis for the Social Sciences and Humanities*. Addison-Wesley.
- Hornak, J. P. (2011). *The Basics of MRI*. Accessed on March 10, 2012.
<https://www.cis.rit.edu/htbooks/mri/images/head.gif>
- Huber, W. 2010. *Kingdom Hearts game play visualizations*. Accessed May 25, 2012.
<http://lab.softwarestudies.com/2012/02/kingdom-hearts-game-play-visualizations.html>
- Husserl, E. 1963. *Ideas: A General Introduction to Pure Phenomenology*. Trans. W. R. Boyce Gibson. New York: Collier Books.
- Inwood, M. J. 2000. *Heidegger: A Very Short Introduction*. Oxford University Press.

- Jenkins, H., Purushotma, R., Clinton, K., Weigel, M., Robison, A. 2006. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. White Paper. Accessed on April 12, 2012.
<http://www.newmedialiteracies.org/files/working/NMLWhitePaper.pdf>
- Kang, Y., Gorg, C., and Stasko, J. 2009. The evaluation of visual analytics systems for investigative analysis: Deriving design principles from a case study. *IEEE VAST*, p.139–146.
- Kantardzic, M. 2003. *Data Mining: Concepts, Models, Methods, and Algorithms*. John Wiley & Sons.
- Kaushik, A. 2007. *Web Analytics: An Hour a Day*. Wiley and Sons.
- Keim, D., Andrienko, G., Fekete, J.D., Gorg, C., Kohlhammer, J. and Melancon, G. 2008. Visual Analytics: Definition, Process, and Challenges. *Information Visualization*. p.154- 175. Springer.
- Kelly, K. and Wolf, G. 2007. Quantified Self. Accessed on April 15, 2012.
<http://quantifiedself.com/>
- Krippendorff, K. 1980. *Content Analysis: An Introduction to Its Methodology*. Newbury Park, CA: Sage.
- Krush DarkGod (alias) and Urme TheLegend (alias). 2009. Darkfall Political Map – Server US1. Accessed on December 5, 2010,
<http://darkfallinfo.com/pmap/?mapserver=us1>.
- Kuper, G. H. and Sterken, E. 2007. Modelling the Development of World Records in Running. In: Albert, J. and Koning, R. H. (eds.) *Statistical Thinking in Sports*. CRC Press, p.7-31.
- Lazzaro, N. 2009. Understand Emotions. In C. Bateman (Eds.), *Beyond Game Design* (pp.3-48). Boston: Course Technology.
- Lee, G. 2010. Halo Wars Stat Tracking Gets a Reprieve. Shack News (online). Accessed on April 12, 2012. <http://www.shacknews.com/article/66760/halo-wars-stat-tracking-gets>

- Lee, H., Smeaton, A., O'Connor, N., Jones, G., Blighe, M., Byrne, D., Doherty, A., Gurrin, C. 2008. Constructing a SenseCam visual diary as a media process. *Multimedia Systems Journal*, 14(6). p. 341-349.
- Liu, Z., Stasko, J., and Sullivan, T. 2009. *Selltrend: Inter-Attribute Visual Analysis of Temporal Transaction Data*. In Proceedings of the IEEE Conference on Information Visualization 2009, p.1025-1032.
- Livnat, Y., Agutter, J., Shaun, M., & Foresti, S. 2005. Visual correlation for situational awareness. *IEEE Symposium on Information Visualization*. p.95-102.
- Lowood, H. 2010. *Players are Artists, too*. Art History of Games, Atlanta.
- MacEachren, A., Boscoe, F., Haug, D. Pickle, L. 1998. *Geographic visualization: designing manipulable maps for exploring temporally varying georeferenced statistics*. Proceedings of IEEE Information Visualization Symposium.
- Macklin, C., Wargaski, J., Edwards, M. and Yang Li, K. 2009. *DATAPLAY: Mapping Game Mechanics to Traditional Data Visualization*. In Proceedings of DiGRA 2009, London, UK, September 1-4: 1-7.
- Magerko, B., Heeter, C., Medler, B. 2010. Tapping into the Hidden Potential of Serious Games: Accommodating Individual Differences. *Interdisciplinary Models and Tools for Serious Games: Emerging Concepts and Future Directions*, ed. Van Eck, R., IGI Global.
- Manovich, L. 2000. Database as a genre of new media. *AI and Society*, 14(2).
- Manovich, L. 2010. *What is Visualization?* Accessed on April 15, 2012.
http://manovich.net/blog/wp-content/uploads/2010/10/manovich_visualization_2010.doc
- Maxis. 2009a. Spore API. Accessed on July 24, 2010, www.spore.com/comm/developer.
- Maxis. 2009b. The Exchange – The Sims 3. Accessed on February 15, 2011, <http://www.thesims3.com/exchange>.

- McCandless, D. 2010. *Information is Beautiful hacks in India with David Cameron*. The Guardian. Accessed on January 2, 2011, <http://www.guardian.co.uk/news/datablog/2010/jul/30/information-beautiful-india-cameron>.
- McElroy, G. 2010. Microsoft closing Halo Wars website, folding into Waypoint community. Joystiq (online). Accessed on April 12, 2012. <http://www.joystiq.com/2010/11/22/microsoft-closing-halo-wars-website-folding-into-waypoint-commu/>
- McGrath, J. 2004. *Loving big brother: performance, privacy and surveillance space*. Psychology Press.
- McHale, I. and Davies, S. 2007. "Statistical Analysis of the Effectiveness of the FiFA World Rankings." In: Albert, J. and Koning, R. H. (eds.) *Statistical Thinking in Sports*. CRC Press, p.77-90.
- Mitchell, R. 2012. Online Pass game included in latest EA server shutdowns. Joystiq (online). Accessed April 12, 2012. <http://www.joystiq.com/2012/03/19/online-pass-game-included-in-latest-ea-server-shutdowns/>
- Medler, B. 2009a. FIFA Earth and Beyond. Accessed April 12, 2012 <http://lcc.gatech.edu/~bmedler3/?p=553>
- Medler, B. 2009b. Generations of Game Analytics, Achievements and High Scores. *Eludamos - Journal for Computer Game Culture*, 3(2), p.177-194.
- Medler, B. 2011. Player Dossiers: Analyzing Gameplay Data as a Reward. *Game Studies Journal*, 11(1).
- Medler, B. and Magerko, B. 2011. Analytics of Play: Using Information Visualization and Gameplay Practices for Visualizing Video Game Data. *Parsons Journal for Information Mapping*, 3(1).
- Medler, B, John, M. and Lane, J. 2011. *Data Cracker: Developing a Visual Game Analytic Tool for Analyzing Online Gameplay*. Paper to be presented at the Conference on Human Factors in Computing Systems (CHI) 2011, Vancouver, BC (Canada).

- Merleau-Ponty, M. 1945. *Phenomenology of Perception*. Taylor & Francis Books Ltd.
- Meyers, A. 2009. Spore Skeletons. Retrieved on January 20, 2010, <http://www.universaloscillation.com/sporeskeletons/>.
- Minecraft X-Ray. 2010. Accessed on April 12, 2012. <http://www.apocalypotech.com/minecraft/xray/>
- Mosteller, F. 1952. The World Series Competition. *Journal of the American Statistical Association*, 47(259), p.355-380.
- MPQ. 2012. In Wikipedia, The Free Encyclopedia. Accessed May 22, 2012. <http://en.wikipedia.org/w/index.php?title=MPQ&oldid=493435402>
- Munsey, P. 2009. *Long arm of law reaches into World of Warcraft*. Kokomo Perspective. Accessed on January 1, 2010. http://kokomoperspective.com/news/local_news/article_15a0a546-f574-11de-ab22-001cc4c03286.html
- Murray, J. 1997. *Hamlet on the holodeck: the future of narrative in cyberspace*. Simon and Schuster.
- Nacke, L., Drachen, A., Kuikkaniemi, K., Niesenhaus, J., Korhonen, H., Hoogen, W., Poels, K., IJsselsteijn, W., and Kort, Y. 2009. *Playability and Player Experience Research*. In Proceedings of DiGRA 2009. London, UK.
- Nicholson, B. 2010. Halo Wars Stat-Tracking, Official Website Gone Dec. 15. Giant Bomb (online). Accessed on April 12, 2012. <http://www.giantbomb.com/news/halo-wars-stat-tracking-official-website-gone-dec-15/2742/>
- Nike, Inc. 2010. Nike+. Accessed on April 15, 2012. <http://nikeplus.nike.com/plus/>
- Norman, D. 1994. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Addison-Wesley.
- O'Reilly, T. 2005. *What is Web 2.0*. Accessed May 25, 2012. <http://oreilly.com/web2/archive/what-is-web-20.html>

- Orszag, P. 2009. Democratizing Data. Accessed on January 2, 2011, <http://www.whitehouse.gov/blog/Democratizing-Data>.
- Perri 6. 1998. *The Future of Privacy vol. 1*. London: Demos.
- Pousman, Z., Stasko, J., and Mateas, M. 2007. Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), p.1145-1152.
- Putnam, R. D. 2000. *Bowling Alone*. New York: Simon & Schuster.
- Radoff, J. 2006. gamerDNA. Retrieved August 21, 2010. www.gamerdna.com
- Rockstar Games. 2008. Rockstar Games Social Club. Accessed on July 24, 2010, <http://socialclub.rockstargames.com/>.
- Rollings, A. and Morris, D. 2004. *Game Architecture and Design: A New Edition*. New Riders.
- Rosling, H. 2005. Gapminder. Accessed on April 15, 2012. <http://www.gapminder.org/>
- S2 Games. 2010. Heroes of Newerth API. Accessed on February 15, 2011, <http://xml.heroesofnewerth.com/>.
- Salen, K. and Zimmerman, E. 2004. *Rules of play: Game Design Fundamentals*. Massachusetts: MIT Press.
- Saraiya, P., North, C., Duca, K. 2005. An Insight-Based Methodology for Evaluating Bioinformatics Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4).
- Schmidt, K. and Pohflepp, S. 2009. Social Collider. Accessed on April 15, 2012. http://postspectacular.com/process/20090320_socialcollider
- Schneider, S. C. 1987. Information overload: Causes and consequences. *Human Systems Management* 7, p.143–153.

- Schneidewind, N., and Hinchey, M. 2008. *Why Predicting Outliers in Software is a Good Thing to Do!* IEEE International Conference on Engineering of Complex Computer Systems. p.91-97.
- Schultze, U., and Vandenbosch, B. 1998. Information overload in a groupware environment: Now you see it, now you don't. *Journal of Organizational Computing and Electronic Commerce* 8(2):127–148.
- Segel, E. and Heer, J. 2010. Narrative Visualization: Telling Stories with Data. *Visualization and Computer Graphics*. 16(6). p.1139-1148.
- Shannon, C. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27 (3), p. 379–423.
- Shedroff, N. 2001. An Overview of Understanding. In *Information Anxiety*, author: Wurman, R. p.27-29.
- Shneiderman, B. 1996. The eyes have it: A task by data type taxonomy for information visualizations. *IEEE Visual Languages*. p.336– 343.
- Shirky, C. 2008. *It's Not Information Overload. It's Filter Failure*. Web 2.0 Expo.
- Schrank, B. 2010. *Play Beyond Flow: A Theory of Avant-Garde Videogames*. Doctoral dissertation. Georgia Institute of Technology, Atlanta.
- Simpson, C. W., and Prusak, L. 1995. Troubles with information overload—Moving from quantity to quality in information provision. *International Journal of Information Management*, 15, p. 413–425.
- Solove, D. 2008. *Understanding Privacy*. Harvard University Press
- Solove, D. 2007. *The future of reputation: gossip, rumor, and privacy on the Internet*. Yale University Press.
- Spence, R. 2000. *Information Visualization*. Addison Wesley.
- Stalh, B. C. 2008. *Information Systems: Critical Perspectives*. Taylor & Francis.

- Star, S. and Griesemer, J. 1989. Institutional Ecology, Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*. 19(3), p. 387-420.
- Stasko, J., Miller, T., Pousman, Z., Plaue, C., Ullah, O. 2004. *Personalized Peripheral Information Awareness through Information Art*. In proceedings of Ubicomp 2004.
- Steam ID Finder. 2010. Accessed on October 28, 2011, <http://steamidfinder.com/>
- Taylor, T.L. 2012. *Raising the Stakes: E-Sports and the Professionalization of Computer Gaming*. MIT Press.
- Terraria Map Viewer. 2011. Accessed May 28, 2011. <http://terrariaworldviewer.codeplex.com/>
- Thomas, J. and Cook, K. 2005. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society.
- Toma, C. L. 2010. *Affirming the Self through Online Profiles: Beneficial Effects of Social Networking Sites*. In Proceeding of CHI 2010, p.1749-1752.
- Tufte, E. 1983. *The Visual Display of Quantitative Information*. Graphics Press.
- Tushman, M. L., and Nadler, D. A. 1978. Information processing as an integrating concept in organizational design. *Academy of Management Review* 3, p.613-625.
- Twardos, M. 2009. Spore Data Visualization [Online]. Accessed on June 10, 2009, <http://thevioletpiece.com/spore/>.
- Valve Corp. 2004. Counter Strike Source. Accessed on February 21, 2011, <http://store.steampowered.com/css>.
- Van Manen, M. 1977. Linking Ways of Knowing with Ways of Being Practical. *Curriculum Inquiry*, 6(3), p.205-228.
- van Wijk, J. 2005. *The Value of Visualization*. In Proceedings of the IEEE Visualization Conference, ed. Silva, C., Groeller, E., Rushmeier, H., p. 85.

- Viegas, F., Perry, E., Howe, E., Donath, J. 2004. *Artifacts of the Presence Era: Using Information Visualization to Create an Evocative Souvenir*. Proceedings of Information Visualization 2004. p.105-111.
- Viegas, F. and Wattenberg, M. 2007. *Artistic Data Visualization: Beyond Visual Analytics*. Proceedings of the Second International Conference on Online Communities and Social Computing. p.182-191.
- Ware, C. 2000. *Information Visualization: Perception for Design*. Morgan Kaufmann.
- Westin, A. 1967. *Privacy and Freedom*, New York: Atheneum.
- Williams, D. & Xiong, L. 2008. Herding Cats Online: Challenges in Deriving a Sample from Online Communities. In Hargittai, E. (Ed.). *Research Methods from the Trenches: The Nitty-Gritty of Empirical Social Science Research*. Ann Arbor, Michigan: University of Michigan Press.
- Wilson, D. 2011. Intentionally Broken Game Design and the Art of “Deputizing” Players. GDC Europe 2011. Accessed on April 12, 2012.
<http://gdcvault.com/play/1014909/Intentionally-Broken-Game-Design-and>
- Winner, L. 1986. *The whale and the reactor: a search for limits in an age of high technology*. University of Chicago Press.
- Wollari (alias). 2008. Dotlan: Eve Maps. Accessed on February 15, 2011,
<http://evemaps.dotlan.net/>.
- Wowhead. 2006. Accessed on June 5, 2011. <http://www.wowhead.com/>.
- Yee, N. 2006. Motivations of Play in Online Games. *CyberPsychology and Behavior*, 9(6), p.772-775.
- Zoeller G. 2010. *Development Telemetry in Video Games Projects*. Presented at the Game Developers Conference 2010.